《 第2回　言語と人文学研究会報告 》

Do you speak, TOEIC?

Aaron Francis Ward

Abstract

This paper will outline a relationship between students' in-class English speaking performance and standardized measures of English reading and listening competence, as measured by the TOEIC exam. Further discussion will also be made of this exam and oral communicative competence in relation to other variables, such as self-reports of general talkativeness, English study at cram-school and time spent living or studying abroad. The results of the study suggest that English oral communicative competence in the forms of fluency and accuracy, is a particularly good predictor of TOEIC listening scores, which are, in turn, negatively predicted by confidence. In contrast, there appears to be no significant relationship between the measures used and student self-reports of talkativeness; who and how often they are inclined to speak to other people in English outside class; nor studying English at a cram school during high school. However, the significant observations that were made regarding TOEIC need to be viewed in the light of other results that indicate that students who have spent greater periods of time studying or living abroad also tend to perform significantly better on TOEIC and on the other measures taken. Finally, discussion is made of the interpretative validity of TOEIC scores as proxies for language competencies that the test does not specifically measure in light of the evidence found.

Introduction

The aim of the current study is to build on ideas developed in a pilot study by Ward（2017）, which preliminarily investigated the relationship between student TOEIC scores and the extent to which students engage with English language outside of class. For the purposes of the current research, the validity of using the TOEIC test as an index of the capacity for students to effectively communicate using oral English language skills, which Bagaric and Djigunovic（2007）describe as communicative competency, will be further explored.

The TOEIC exam is perhaps the most widely used standardized measure of English language proficiency in Japan, and was first implemented on a broad scale in 1979. Its use has peaked in recent years with over two-and-a half million candidates taking the examination per year（The Institute for International Business Communication 2019）. TOEIC is administered to recruit and stratify students at all tiers of the education sector and is also used as an instrument for organizing the labour market. Although used as a global index of language competence, the most commonly used variant of the test consists only of multiple-choice listening and reading sections, which constrain the instrument's validity as a measure of dialogic communicative competence. The necessity for TOEIC to be delivered on a mass-scale also limits the kind of measures that can be directly inferred from test scores about candidates' productive competences. Nonetheless, TOEIC scores are routinely used to sort candidates according to their oral communication and writing skills. TOEIC also has broad reaching implications for graduates in the labour market, where employees are often required to undertake the test as part of recruitment and workplace initiation processes. In 2018 approximately 48, 000 job applicants, as well as approximately one million staff had to take the test as part of professional evaluation（The Institute for International Business Communication 2019）. The importance of TOEIC in Japan's educational and labour sectors has potentially also led to an unusually large emphasis on test performance over language acquisition, *per se*.

This is indexed by the existence of a market for instruction and study materials geared towards test-taking tactics, as evidenced by books such as, *Tactics for TOEIC*（Trew 2008）and *Mark Your Goal*（Stafford, Nakata, and Mizumoto 2009）, which emphasize strategies for answering test items based on question-format logic, rather than language knowledge.

Separate speaking and writing components have recently been developed for TOEIC, and while the use of these assessment tools has been increasing, uptake has been relatively slow. The number of candidates for these sections of the test in 2018 accounted for only 1.6％ of total test takers that year. That said, data released by the exam administrators indicate that TOEIC speaking-test performance appears to increase relatively linearly with combined listening and reading test scores. However, data from the same set show an unusual relationship between the four skills measured, in that candidates tended to perform noticeably better in the listening section compared to the reading, but better in writing than speaking（The Institute for International Business Communication 2019）. Further, the new speaking test has, like the listening and reading sections, been designed for mass-administration, and is done online, without a present nor active examiner. Rather, each candidate's test answers are "digitally recorded and sent to ETS's online Network for Evaluation where they are scored by certified ETS raters."（ETS.org 2019, 24）. In contrast, other popular English proficiency exams, such as IELTS and Cambridge suite examinations, use an on-site examiner and/or speaking-test partner, which allows examinees to draw on paralinguistic communication skills such as body-language and back-channeling（c.f. Bagaric and Djigunovic 2007）.

In sum, the relative lack of emphasis placed on speaking skills in TOEIC, in tandem with its widespread use may mean that candidates or institutions who institute its use may lack valid indices of communicative competence. Further, the restrictive nature of the TOEIC speaking test may mean that little is known about candidates actual ability to deploy, in-context, the immediately demanding and reflexive language skills re-

quired of oral communication.

Research Questions

1. How useful are TOEIC listening and reading scores as a measures of in-class student speaking competence?
2. What factors outside the classroom may also influence student in-class speaking competence?

Method

The current research was conducted in two phases. All of the participants were first-year university students, who were all enrolled in a compulsory, functional-language based, unified English communication program, that did not assess students' attendant accuracy or content. As such, all of the participants, regardless of their TOEIC-based assignment to English proficiency-level classes, were subject to identical learning expectations and outcome measures. Further, none of the measures used in the current study were part of the course's assessment nor had any discernable impact on lesson flow. This gave the research an ethological nature, in that the participants were unable to perform to this study's variables of interest on the expectation that doing so would have any positive or negative effect on their grades.

Phase 1

For the first phase of this research, a subset of the participants was observed twice for 90 minutes midway through and at the end of their course of study. In each lesson, the participants were required to spend at least 50 minutes engaging in non-task-based pair-work and small group discussions, without any teacher intervention. Each class had no more than nine students, making relatively close observation of each participant possible. During observation lessons, each participant was scored from 1（lowest）to 5（highest）on the following criteria: *fluency*, *accuracy*, *confidence*, *willingness to use*

English and *attention to task*, according to the criteria in Appendix I. In addition to the variables of interest, field notes were taken during the observations to help qualify the data collected.

Phase II

In Phase II, the participants were given a questionnaire, in Japanese, regarding their engagement with English outside the classroom. This included: self-reports of general talkativeness, frequency of spoken English use outside class, English study at cram school during high school, and length of time spent studying or living abroad. This data was collated at the end of the semester to reduce rater bias in Phase I. For further detail, please refer to Appendix II.

Results

Sample

A total of 576 first-year students from a private university in Tokyo participated in this study from 2016 to 2018. However, not all of the students participated in both phases of the study. Although the majority of the participants were Japanese（95.1％）, the non-Japanese respondents have been removed from the sample, since no one other ethnic group exceeded the 5％ margin of error used in this report, leaving 536 participants. Of this sample, 311（58.3％）were female and 225（41.7％）were male.

Table 1 Participant TOEIC Listening and Reading Scores

	Sample Means	
Gender	Listening	Reading
Male	253.26	221.38
Female	299.59	249.06
$t_{(527)}$	-4.91**	-3.11**

** $p < 0.01$（two-tailed, equal variances assumed
n = male（225）, female（311）

The results in Table 1 indicate that female participants scored significantly higher on TOEIC listening and reading than male participants.

Table 2 Correlations between TOEIC Listening and Reading Scores and Variables of Interest

TOEIC r^2	n	Fluency	Accuracy	Confidence	Willingness	Attention
Listening	369	.644**	.733**	.320**	.206**	.306**
Reading	369	.517**	.572**	.191**	.178**	.298**

** $p < 0.01$

Table 2 shows that all of the variables of interest were positively correlated with the participants' TOEIC scores.

Table 3 Comparison of Male and Female Participants on Variables of Interest

Male - Female	Fluency	Accuracy	Confidence	Willingness	Attention
t-score	-4.607**	-4.134**	-1.033 (ns)	-0.690 (ns)	-2.677 (ns)

** $p < 0.01$, equal variances assumed, n = male (154) , female (216)

Table 3 indicates that, in line with their TOEIC scores, female participants scored more highly than males on the variables of interest, although only fluency and accuracy were significantly different.

Table 4 Model Fits for Variables of Interest as Predictors of TOEIC Scores

	F-Score	R^2
TOEIC Listening	$F_{(5,363)} = 93.03$**	0.56
TOEIC Reading	$F_{(5,363)} = 43.18$**	0.37

** $p < 0.01$

Table 4 indicates that the models generated by the linear regression analyses are both significant. The variables of interest predict 56 % of the total variation of the participants' listening scores. While the model also significantly predicts the participants' reading scores, this only accounts for 37% of the total variation observed and produces

a notably lower F-score than for listening. Therefore, TOEIC listening scores appear to be a better index of student in-class speaking competence than reading scores.

Table 5 Linear Regression of Variables of Interest and Participant TOEIC Scores

	Standardized Coefficients			
	TOEIC Listening		TOEIC Reading	
	β	t	β	t
Constant		2.81**		2.80**
Fluency	.202	2.89**	.260	3.10**
Accuracy	.688	10.44**	.478	6.07**
Confidence	-.152	-3.12**	-.283	-4.88**
Willingness	-.039	-.985（ns）	-.051	-1.08（ns）
Attention	-.053	-1.14（ns）	.095	1.69（ns）

** $p < 0.01$

Tables 5 shows that there is a significant constant factor in both models, that fluency and accuracy are significant and positive predictors of TOEIC scores, and that confidence is a significant but negative predictor. Once regressed, willingness-to-speak-English and attention-to-task cease to be significantly related to TOEIC.

Table 6 Participant Self-Reports of Talkativeness

	Avoid Talking	Close People	Enjoy Talking	New People	Love Talking *
Female	2.2%	11.0%	35.4%	34.3%	17.1%
Male	7.8%	12.5%	37.5%	31.3%	10.9%
Total	4.5%	11.7%	36.2%	33.0%	14.6%

* For self-rating descriptors, please refer to Appendix II.

A Chi-squared analysis（$\chi^2_{(4, 309)} = 7.60$, $p = 0.12$（ns））of the results in Table 6 indicate that, although the participants generally report themselves to enjoy talking to other people, and that females report themselves to be more talkative, these results are statistically independent of one-another. As such, unlike the Phase I variables of interest, gender did not have any significant effect on self-reported talkativeness.

Table 7 Participant Self-Reports of Frequency of English Spoken outside Class by Gender

			Frequency						
	Gender	n	never	less than once a month	about once a month	about once a week	several times a week	every day	χ^2
Speaking to Japanese Friends	Female	291	57.0%	4.8%	4.1%	5.2%	9.3%	19.6%	3.26 (ns)
	Male	208	61.1%	4.3%	1.9%	5.3%	6.7%	20.7%	
Speaking to non-Japanese Friends	Female	290	63.4%	9.3%	7.6%	6.9%	7.9%	4.8%	1.97 (ns)
	Male	206	66.5%	8.7%	6.3%	8.3%	5.3%	4.9%	
Speaking to non-Japanese Professors	Female	291	55.0%	6.5%	5.5%	15.5%	11.7%	5.8%	7.96 (ns)
	Male	206	56.3%	6.8%	3.4%	18.4%	13.6%	1.5%	
Speaking to Strangers	Female	290	51.4%	19.7%	12.8%	6.6%	8.3%	1.4%	0.12 (ns)
	Male	205	59.0%	19.5%	8.8%	3.9%	6.8%	2.0%	

* For self-report descriptors, please refer to Appendix II.

The results in Table 7 indicate that the results are statistically independent of one-another, meaning that again gender did not have a significant effect on speaking English outside class.

Table 8 Variables of Interest by Cram School Attendance

Variable	t-test (attend — not attend)*
TOEIC Listening	$t_{(480)} = -1.38$, (ns)
TOEIC Reading	$t_{(480)} = -0.72$, (ns)
Fluency	$t_{(349)} = -0.88$, (ns)
Accuracy	$t_{(349)} = -1.70$, (ns)
Confidence	$t_{(349)} = -0.64$, (ns)
Willingness	$t_{(349)} = -0.08$, (ns)
Attention	$t_{(349)} = 0.02$, (ns)

* equal variances assumed

The results in Table 8 indicate that having studied English at a cram school during high school tended to have negative but non-statistically significant effects on all of the

variables of interest, except *attention to task*.

Table 9 Variables of Interest by Time Spent Living or Studying Abroad

Variable	F-test
TOEIC Listening	$F_{(3,489)} = 46.23^{**}$
TOEIC Reading	$F_{(3,489)} = 85.49^{**}$
Fluency	$F_{(3,351)} = 27.23^{**}$
Accuracy	$F_{(3,351)} = 36.93^{**}$
Confidence	$F_{(3,351)} = 9.77^{**}$
Willingness	$F_{(3,351)} = 1.85$ (ns)
Attention	$F_{(3,351)} = 4.83^{**}$

** $p < 0.01$

The results in Table 9 indicate that living or studying abroad had a significant positive effect on all of the variables of interest except *willingness to use English*. Post-hoc Tukey analyses show a general progression of higher scores on the variables of interest the longer the participants had spent abroad. These homogeneous subsets most often occurred between the participants who had spent six months or longer abroad, and those who has spent less time or had not been abroad at all. Although there was a significant difference between the study abroad groupings for *attention-to-task*, but not *willingness-to-use-English*, there were no homogeneous subsets for either of these variables.

Table 10 Participant Self-Reports of Frequency of English Spoken outside Class by Length of Time Abroad

			Frequency						
	Length	n	never	less than once a month	about once a month	about once a week	several times a week	every day	χ^2
Speaking to Japanese Friends	Never	286	67.5%	3.5%	1.7%	5.2%	5.2%	16.8%	57.49** †
	0 − 1 month	101	62.4%	5.0%	4.0%	5.9%	5.9%	16.8%	
	1 − 3 months	18	38.9%	11.1%	5.6%	0.0%	16.7%	27.8%	
	6 − 12 months	23	21.7%	8.7%	8.7%	17.4%	17.4%	26.1%	
	1 year +	66	37.9%	4.5%	6.1%	1.5%	18.2%	31.8%	
Speaking to non-Japanese Friends	Never	284	80.3%	7.7%	3.5%	4.9%	2.5%	1.1%	149.99**
	0 − 1 month	102	56.9%	15.7%	9.8%	6.9%	5.9%	4.9%	
	1 − 3 months	17	47.1%	0.0%	23.5%	5.9%	23.5%	0.0%	
	6 − 12 months	23	30.4%	17.4%	8.7%	26.1%	13.0%	4.3%	
	1 year +	65	30.8%	3.1%	12.3%	12.3%	21.5%	20.0%	
Speaking to non-Japanese Professors	Never	285	64.9%	5.6%	3.2%	17.2%	7.0%	2.1%	91.06** †
	0 − 1 month	102	57.8%	6.9%	6.9%	13.7%	9.8%	57.8%	
	1 − 3 months	17	52.9%	17.6%	5.9%	11.8%	11.8%	0.0%	
	6 − 12 months	23	8.7%	26.1%	4.3%	26.1%	30.4%	4.3%	
	1 year +	65	32.3%	1.5%	7.7%	15.4%	30.8%	12.3%	
Speaking to Strangers	Never	285	63.5%	16.8%	9.8%	3.2%	6.0%	0.7%	59.54** †
	0 − 1 month	102	52.9%	23.5%	6.9%	4.9%	10.8%	1.0%	
	1 − 3 months	17	41.2%	11.8%	23.5%	5.9%	5.9%	11.8%	
	6 − 12 months	23	30.4%	26.1%	13.0%	17.4%	13.0%	0.0%	
	1 year +	64	32.8%	21.9%	20.3%	12.5%	7.8%	4.7%	

** $p < 0.01$, † cell count assumption violation

Table 10 indicates that a larger percentage of the participants reported that they never speak English outside class to anyone. However, in each categorization of persons-spoken-to, there is a tendency for this percentage to decrease the longer the participant has spent living or studying abroad. However, as the frequency of reported spoken interactions increases, this pattern becomes less clear. Further, although Chi-

squared analyses suggested that these results are not statistically independent of one-another, in all but the *speaking to non-Japanese friends* categorization, there were violations of the minimum cell-count assumption of the Chi-squared test. This is a result of too few participants having lived or studied abroad for 1 − 3 and 6 − 12 months. As such, having spent more time abroad had a significant effect on how frequently participants talk to non-Japanese friends in English, but more data is needed to clarify the other categorizations.

Table 11 Participant Self-Reports of Rationale for English Spoken outside Class by Length of Time Abroad

			Rationale for Speaking English outside Classroom				
		n	Avoid	Indifferent	Sometimes Practice	Good Practice	Main Reason *
Length of Time Abroad	Never	155	65.2%	14.2%	8.4%	7.1%	5.2%
	0 − 1 month	46	50.0%	13.0%	23.9%	4.3%	8.7%
	1 − 3 months	15	40.0%	26.7%	33.3%	0.0%	0.0%
	6 − 12 months	20	15.0%	15.0%	30.0%	25.0%	15.0%
	1 year ＋	49	14.3%	32.7%	18.4%	14.3%	20.4%

$\chi^{2}_{(16,16)} = 69.59$, $p < 0.01$, cell count assumption violation

* Refer to Appendix II for full descriptors.

The results from Table 11 suggest the longer the participants had spent studying or living abroad the more likely they were to rationalize speaking English outside of class as a means to improving their English language competence. The Chi-squared analysis suggest that these results are not significantly independent of one-another. However, this result must again be viewed tentatively due to a cell-count assumption violation.

Discussion

The results of this study indicate that while speaking skills are predictive of TOEIC, the results may not index a direct relationship between TOEIC test performance and communicative competence. Firstly, the female participants in the study, on average,

significantly outperformed their male counterparts on the TOEIC exam, as well as on a number of other observations made in Phase II. This lends some measure of external validity to the instruments used in Phase I in that gender has consistent effects on the two sets of measures.

While the variables of interest in Phase I are all significantly and positively correlated with TOEIC, the strength of these relationships varies widely from 0.733（listening & *accuracy*）to 0.178（reading & *willingness to communicate*）. Further, when directly compared, the female participants also significantly outperformed the males in fluency and accuracy. When all of these variables were linearly regressed against one-another, both *fluency* and *accuracy* continued to be both positively and significantly predictive of variation in TOEIC scores, and *confidence* became significantly negatively predictive. However, *willingness to communicate in English* and *attention to task* ceased to be statistically significant altogether. These regression analyses show that the model accounts for more variation in TOEIC listening scores than reading. The former also generated a noticeably larger F-score. This somewhat contradicts the pattern of test results published by the TOEIC test administrators（The Institute for International Business Communication 2019）, in that TOEIC listening scores appear to be a better proxy for spoken English fluency and accuracy than reading scores. Nonetheless, the results of this study are consistent with Feingold's（1994）meta-analytical studies of gender and personality, which indicate that males tend to be more assertive than females, who in turn are more conscientious and agreeable. Further, studies into gender and language use, as summarized by Xia（2013）, show that women tend to use language more in line with the kinds of elaborated code required of academic contexts, which are often found in written form and testing materials（Bernstein 2003）. This existing research helps to contextualize why female TOEIC candidates tend to outperform males, and why they appear to use language more contentiously in class, despite confidence contributing negatively to performance.

It must be kept in mind that the results of Phase I indicate covariation between TOE-

IC and the variables of interest, not causality. The results of Phase II help to qualify the relationships observed. Firstly, the majority of the participants reported that they enjoy talking to other people they know and/or enjoy talking to different kinds of people, including new people. Although a larger percentage of the male students reported that they avoid talking to others where possible, there was no significant difference between the male and female participants across the self-categorizations. This indicates that self-reported talkativeness does not explain the results of Phase I. Further, when asked how often the participants spoke to people in English outside of class, depending on who was being spoken to, between 51.4％ and 66.5％ of the participants reported that they never do this at all. Again, there was no significant difference between the male and female participants in this regard, suggesting that the factors behind these patterns differ from those of Phase I（c.f. Feingold 1994）. There was, however, a subset of participants who appeared more likely to talk to Japanese friends every day and non-Japanese professors outside class several times a week. This result could be a consequence of some participants making an extra effort to do so, or because they enrolled themselves in other classes that required it. Either way, through either study choices or otherwise, these participants actively engage outside of class in English more often than the others.

To qualify the results presented so far, the current research will now turn to factors outside the immediate educational setting that may influence the participants' oral communicative competence. One Japanese extra-curricular setting which may influence language learning outcomes is attendance at cram schools, which are particularly geared towards helping students with their grades and university entrance examinations（Tofugu 2019）. However, the results of this study suggest that cram school did not have any significant effects on either TOEIC scores or the variables of interest in Phase I. In fact, cram school attendance mostly seemed to have a negative influence on performance. This reinforces the prospect that TOEIC performance proper is not the sole driver of communicative competence.

The results presented so far are somewhat vexing. While many of the participants appear to be quite communicatively competent in both English and Japanese, and this is related to their TOEIC scores, it does not seem to be the straightforward consequence of either talking to people outside class or having taken additional English language studies in the form of cram school. However, one factor which seems to best delineate the participants in these regards is the amount of time they had spent studying or living abroad. The results of this study indicate that the longer the participants had spent living abroad, the significantly better they performed on the TOEIC test, as well as all of the variables of interest in Phase I, excluding *willingness to speak English* in class. The amount of time spent abroad seemed to have the greatest effect on *accuracy* and *fluency* respectively, and unlike when regressed against TOEIC scores, has a positive effect on confidence. Further, post-hoc analyses tended to indicate that significant differences between the groups of participants were mostly attributable to students who had spent at least six months abroad. Additionally, there seems to be an inverse relationship between the length of time spent abroad and the number of participants reporting that they never speak in English outside class - particularly to Japanese and non-Japanese friends. However, as the frequency of reported speaking English outside class increases, these patterns become less clear. These results need to be treated cautiously, due to the low number of participants who had spent between 1 − 3 and 6 − 12 months abroad. Finally, the participants who had spent longer abroad generally had noticeably more positive attitudes about the value of using English outside class to those who had never been abroad or had only been for relatively short lengths of time. While it is arguably unclear whether the earlier-reported tendency for participants to overwhelming avoid speaking English outside class was a consequence of a lack of interest or opportunity. However, when recast in terms of the length of time spend abroad, there is evidence that the participants do have some degree of opportunity to speak English if they are so inclined to do so. This also suggests that the results for cram school may need to be qualified by further research. It could well be that, for example, the partici-

pants who spent significant periods of time abroad did not attend cram school. Consequently, better performance on TOEIC by students who had lived abroad for longer stays may well have ironed out any of the benefits of cram school, as observed in this study.

The results of this study tend to indicate that the TOEIC scores used in Japanese universities to allocate students to English proficiency-level classes actually perform relatively well. However, TOEIC listening scores alone are probably more useful for student allocation to English communication classes. Nonetheless, it seems that TOEIC scores should only be viewed as a proxy for communitive competence in that time spent studying or living abroad is a significant and compounding factor in both TOEIC test performance, as well as oral communicative competence. These results are, of course, not particularly surprising, given the increased opportunity and necessity that living abroad demands of students to use their oral communication skills, in particular.

However, the benefits of living abroad are not the cumulative effects of language exposure alone. They are also a consequence of a degree of enculturation in the target language that is significantly less accessible to students who study a second language in their own language context. According to Kramsch（1998）, language as it is taught in the classroom is presented as relatively discrete linguistic chunks. For practical reasons, this pedagogic strategy is makes sense as a means to scaffolding language acquisition. However, these necessarily artificial learning contexts lend themselves to a nomothetic approximation of how language occurs in actual contexts. As such, living abroad not only increases the exposure of students to the target language and reduces the prospects for avoidance, it most importantly provides learners with a tacit understanding of how the game of language plays out in situ（Wittgenstein 1958）. In sociological terms, life abroad provides students with a *feel for the game*–an invaluable, reflexive understanding of, not just the rules of the language, but also an enculturated understanding of how and when to apply, twist, bend and break them（Bourdieu 1991）. This, in part, provides an understanding of why the results of this study indicated a less

consistent relationship between time spent abroad and *willingness to use English* and *attention to task*. A relatively consistent pattern emerged from the field notes taken during the current study in the students assigned to higher-proficiency level classes who had *apparently* spent some time abroad（i.e. those without a pronounced katakana accent）. These students were noticeably less focused on speaking English during lessons at times when they did not think they were being assessed. This is perhaps not surprising, as they had relatively successfully enculturated themselves into two distinct language contexts, where performing in the classroom context had no effect on their test performance, and in turn stood to socially distance them from their peers.

While it seems that one of the most significant contributors to oral English communication competence comes from being able to spend at least six months studying aboard, this requires significant financial commitment. The analyses of the current study indicate that students who have had less L2 cultural experience also tended to avoid speaking to people in English outside of class. However, one alternative to this may be to encourage students early in the language learning process to foster an interest in the culture of the target language. One strategy often employed to this end is to introduce popular music into the classroom（e.g. Poulshock and Menish 2014）. However, as preliminary data suggest, student self-reported interest in the associated culture of L2 is typically inversely related to the difficulty of engaging with particular aspects of cultural production（Ward 2017, Ward 2018）. As such, strategies which focus on poetic culture, such as listening to popular music in class, may be difficult to adapt to target language, and worse, may act to reinforce students' tacit preconceptions about the relationship between culture and learning. Prosaic culture, on the other hand, can be both more adaptive and could be, long term, more beneficial. No pain, no gain.

Conclusion

The results of the current study demonstrate that the TOEIC test, in spite of generally not assessing spoken English, is actually a relatively good index of oral communica-

tive competence. However, allocation of students to proficiency-based classes would be better done using TOEIC listening scores in tandem with knowledge about how much time students have spent abroad. The current paper is still only a preliminary investigation into the role of culture in language acquisition. Many questions still remain unanswered, and there is great scope for further investigation into how Japanese English-language students engage with or avoid attendant culture, and how this affects their learning and language acquisition. This research could include an examination of the kind of high school attended and the relationship between the kinds of extra-curricular clubs and circles students engage in and their studies.

Appendix I: In-class Analysis Criteria

Criterion	Score	Descriptors
Fluency	1	Speaks with a very noticeable number of pauses and takes significant time before continuing
	2	Frequent pauses and takes significant time before continuing
	3	Intermittent pauses, often takes time before continuing
	4	Occasional pauses, but is able to continue quickly
	5	Continues speaking at a proficient length and speed without pausing
Accuracy	1	Agrammatical, rarely uses function words, and has limited vocabulary which impedes communication
	2	Usually speaks using simple sentence patterns, relies on a limited range of vocabulary and basic verb forms, and often misuses or does not use function words
	3	Usually speaks using simple sentence patterns and verb forms and has a basic mastery of function words and has a sufficient vocabulary to talk about familiar topics
	4	Is able to use complex sentence patterns with sub-clauses, a variety of verb forms, usually uses function words, has a sufficient range of vocabulary to discuss familiar topics
	5	Grammatically accurate spoken English with a wide vocabulary range at or close to a proficient level

Confidence	1	Appears comfortable only during very controlled activities, and appears very anxious during preparation and fluency activities
	2	Appears comfortable doing very controlled and semi-controlled activities, but usually appears anxious during preparation and fluency activities
	3	Appears comfortable doing controlled and semi-controlled practice and preparation activities, but sometimes appears to be anxious during preparation activities and appears to be hesitant about joining fluency activities
	4	Appears comfortable doing warm-up, controlled practice and preparation activities, but sometimes waits to be invited by other students before joining fluency activities
	5	Appears very comfortable communicating with others, particularly in preparation and fluency activities, usually initiates interactions with other students during activities
Willingness to Communicate in English	1	Usually relies on L1 to articulate ideas and/or resolve communication breakdowns, frequently makes slips in L1
	2	Often relies on L1 to articulate complex vocabulary and/or resolve communication breakdowns, often makes slips in L1
	3	Sometimes relies on L1 to articulate complex vocabulary and/or resolve communication breakdowns, sometimes makes slips in L1
	4	Rarely relies on L1 to articulate complex vocabulary and/or resolve communication breakdowns, rarely makes slips in L1
	5	Never relies on L1, even during substantial communication breakdowns
Attention to Task	1	Rarely follows instructions, often distracts others, and uses target language infrequently and expediently
	2	Usually follows instructions , sometimes distracts others, and uses target language intermittently and expediently
	3	Usually follows instructions, occasionally distracts others, sometimes talks off topic and is occasionally distractible, but uses target language appropriately
	4	Often on task, seldom distracts others or talks off topic and is infrequently distractible, and makes a good effort to use target language
	5	Always on task, redirects others back to the task and maximises opportunities to use target language

Appendix II: Survey Battery Items

* Note: Only the Japanese version of this survey was administered.

1. Which statement best describes you about talking to other people outside class?

授業外で人と会話をすることについて。あなたはどれに当てはまりますか？

English	I avoid talking to other people whenever possible.	I like to talk to my close friends, but I would rather not talk to people who I do not know or do not know well.	I enjoy talking to people, but I feel uncomfortable when talking to people who I do not know.	I enjoy talking to different kinds of people, and like to meet new people.	I love to talk to many different kinds of people, and try to meet new people whenever possible.
Japanese Translation	人と話すことをできる限り避ける	仲の良い友達とは話すが、知らない人とはできるだけ話さないようにしている	人と話すことは楽しいが、知らない人と話すのは落ち着かない	様々な人と話すことを楽しく感じ、人と出会うことが好きだ	たくさんの人と話すことが好きで、新しい人とできるだけ出会いたいと思う

2.　Outside class, how often do you speak to the following people in English?

授業外に、どれくらいの頻度で下記の人と英語で会話をしますか？

Japanese friends / 日本人の友達	never	less than once a month	about once a month	about once a week	several times a week	every day
non-Japanese friends / 外国人の友達	話さない	1ヶ月に1回以下	1ヶ月に1回程度	1週間に1回程度	1週間に数回	毎日
non-Japanese teachers/professors / 外国人の教師／教授	never	less than once a month	about once a month	about once a week	several times a week	every day
people who you do not know（customers, people on the street, etc.）/ 他人（客、道端で出会った人等）	話さない	1ヶ月に1回以下	1ヶ月に1回程度	1週間に1回程度	1週間に数回	毎日

3.　Which statement best describes your reasons for speaking English outside class?

授業外で人と英語で会話をすることについて。下記のどれに当てはまりますか？

English	I only talk to other people in English when I absolutely have to.	I'm not particularly concerned if talking to other people in English helps to improve my language skills.	I sometimes talk to people in English to improve my language skills.	I talk to people in English because I think is a good way to improve my language skills.	The main reason I talk to people in English is to improve my language skills.
Japanese Translation	どうしても英語で話さなければならない時に他の人と英語で話します	他の人と英語で話すが、それは英語力を向上させるために行っているのではない	英語力を向上させるために、たまに人と話す	英語力を向上させるためには良い方法だと思うため、他の人と英語で話す	他の人と英語で話す理由は英語力を向上させるためである

4.　How long have you lived or studied abroad for?

海外で生活や学校に通ったことはありますか？

never	for less than one month	for 1−3 months	for 6−12 months	for more than one year
ない	1ヶ月以下	1〜3ヶ月ほど	6から12ヶ月ほど	1年以上

5. When you were a high school student, did you study English at a cram school?
高校生のころに、塾で英語の授業を受けていましたか？

a. Yes / はい

b. No / いいえ

References

Bagaric, Vesna, and Jelena Mihaljevic Djigunovic. 2007. "Definining Communicative Competence." *Metodika* 8（1）: 94−103.

Bernstein, B. 2003. *Class, Codes and Control, Volumes I – IV*. 4 vols. Vol. 1−4. London: Routledge.

Bourdieu, Pierre. 1991. *Language & Symbolic Power*. Cambridge, Massachusetts: Harvard University Press.

ETS.org. 2019. Examinee Handbook: Speaking & Writing. edited by English Testing System.

Feingold, Alan. 1994. "Gender Differences in Personality: A Meta-Analysis." *Psychological Bulletin* 116（3）:429−456.

Kramsch, Claire. 1998. *Language and Culture*. Oxford: Oxford University Press.

Poulshock, Joseph, and Marc Menish. 2014. "Knowing Culture Through Music." *JALT Conference Proceedings* 1:428−436.

Stafford, Mark D, Tatsuya Nakata, and Atsushi Mizumoto. 2009. *Mark Your Goal: Vocabulary and Grammar Tactics for the TOEIC test*. Tokyo: Kinseido.

The Institute for International Business Communication. 2019. TOEIC Program: Data & Analysis 2019. New Jersey: Princeton.

Tofugu. 2019. "Let's Talk about Japanese Cram School." accessed December 1. https://www.tofugu.com/japan/japanese-cram-school/.

Trew, Grant. 2008. *Tactics for TOEIC（R）Speaking and Writing Tests: Pack : Tactics-focused preparation for the TOEIC（R）Speaking and Writing Tests* Oxford: Oxford University Press.

Ward, A.F. 2018. "The C in EFL." *New Directions in Teaching and Learning English Discussion* 6:In press.

Ward, Aaron Francis. 2017. "Culture, Language, Attitude, Performance." *New Directions in Teaching and Learning English Discussion* 5:233－242.

Wittgenstein, Ludwig. 1958. *Philosophical Investigations*. Translated by G. E. M. Anscome. Third ed. Oxford, UK: Basil Blackwell Ltd.

Xia, Xiufang. 2013. "Gender Differences in Using Language." *Theory and Practice in Language Studies* 3（8）:1485−1489.

―Aaron Francis Ward・東洋大学法学部講師―