

KANSAI GAIDAI UNIVERSITY

Accounting for Communicative Competence via Pragmatics : A Pilot Test of Instruments

メタデータ	言語: eng 出版者: 関西外国語大学・関西外国語大学短期大学部 公開日: 2016-09-05 キーワード (Ja): キーワード (En): Pragmatics, Speech Acts, Willingness to Communicate, Rasch modeling, ANOVA 作成者: Rylander, John W., Clark, Phillip M., Derrah, Richard, Shinall, J. Scott メールアドレス: 所属: 関西外国語大学, 関西外国語大学, 啓明学院高等学校, 関西外国語大学
URL	https://doi.org/10.18956/00006155

Accounting for Communicative Competence via Pragmatics

A Pilot Test of Instruments

John W. Rylander, Phillip M. Clark, Richard Derrah,
and J. Scott Shinall

Abstract

This study reports the results of a pilot-testing project of a video-based pragmatics instrument whose purpose was the assessment of the receptive level of communicative competence of high school and university students in the Japanese context. The instrument was analyzed using Rasch modeling and the results show high levels of item reliability with low levels of person reliability for the sample population. Person logits of ability on the pragmatics assessment were subjected to linear regression (ANOVA) using SPSS as a criterion variable against a set of predictor variables including scores on EIKEN, TOEFL, TOIEC, a 64-item Willingness to Communicate motivation survey, educational year, gender, experience learning in an English-medium school, and time spent in residence abroad. Results indicate three variables to be statistically significant predictors of pragmatic ability: TOEFL score, year of education (years 2 and 3 at the high school level and year 1 and 2 in university), and residence abroad.

Keywords: Pragmatics, Speech Acts, Willingness to Communicate, Rasch modeling, ANOVA

Introduction

This study is explorative in nature and serves as pilot research for a future structural equation model (SEM) to assess sets of predictor variables and their relationships to the latent variable *communicative competence*. The operationalization of this variable is viewed as learner performance on a battery of assessments designed to account for knowledge of specific pragmatic features of English. The goal of the SEM will be to confirm the hypothesis that the most efficient acquisition of pragmatics knowledge (i.e., the pragmalinguistic forms and sociopragmatic features) results from explicit instruction rather than implicit input.

The research involved the construction of a multiple-choice pragmatics instrument, the

employment of a motivational survey, the collection of various biographical data, the analysis of each set of data individually, and then a combined analysis using linear regression. Section one will briefly describe the sample population, allowing for a basic understanding of the generalizability of the findings. Section two details the construction, implementation, and analysis of one component of the pragmatics assessment — a video-based (i.e., receptive) instrument of learner awareness and sensitivity to a variety of speech acts (SA) (e.g., *apologies, invitations, offers, requests, suggestions*). The final section provides the analysis of the statistical data — including gender, context (high school vs. university), TOEFL score, TOIEC score, EIKEN score, residence abroad, enrollment in an English-medium institution, and measurements on a multi-construct motivation survey. The assumption is that from this collection of data a set of possible predictor variables would emerge for use with SEM. Expectations regarding results were tempered, and as a methods paper the significance of the findings should be viewed as such.

Participants

The sample initially comprised 377 Japanese student learners of English in intact classrooms. All participants sat both the pragmatics test and motivation survey. These were administered on different days — as a result, absenteeism, lack of ID number entry, and problems with accurately entering responses reduced the total sample size to 346. Of this sample ($n = 346$), 35 were 2nd year high school students, 179 were 3rd year high school students, 55 were 1st year university students, and 77 were 2nd year university students.

The high school students attend a private co-ed high school with an approximate population of 650 students divided into co-ed and all-girls divisions. Subjects were 2nd and 3rd year native-speakers of Japanese enrolled in the co-ed division. The test was administered within separate classes with class sizes averaging 20 students. The co-ed division participates in a recommendation program for entrance into a university in the Kansai area and scores on the EIKEN test are a requirement. During a typical week the students have from 5 to 7 class hours of English instruction; of these, two hours are “Oral Communication” classes conducted by native speakers of English. Most other classes are conducted in Japanese.

The undergraduate students were from a medium-sized private foreign language university (*gaikokugo daigaku*) in the Kansai region and were native speakers of Japanese. Of those students who reported their TOEFL score ($n=119$), the scores ranges between 370-603 (mean=488). The students had all studied English as a Foreign Language in the Japanese

secondary school system for an average of 6-7 years prior to participating in the study, with 50 of these students having lived abroad in an English-speaking country for various lengths of time. Each student has from 6 to 12 hours English instruction weekly from native speakers of English. First year undergraduates were surveyed midway through their 1st semester, after approximately 80 hours of English instruction. Second year university students were surveyed in the 1st semester of their 2nd year, after approximately 430 hours of English instruction.

Creation, Employment and Analysis of the Criterion

A 38-item instrument was designed to assess L2 learners' meta-pragmatic awareness of 10 speech acts (SA): *apologies*, *compliments*, *farewells*, *gift-givings*, *greetings*, *introductions*, *invitations*, *offers*, *requests* and *suggestions*. Figure 1 shows response options, including distractors both represented by other clips and those assumed not present in any clips (highlighted, for present purposes, by their lack of Japanese translations).

Figure 1: Speech Act Sample

Apology: 謝罪	Gift-Giving: 贈り物の贈呈	Offer: 提供、申し出
Command:	Greeting: あいさつ	Permit:
Compliment: 賛辞、ほめ言葉	Guess:	Promise:
Farewell: 別れのあいさつ	Introduction: 紹介	Request: リクエスト
Forgive:	Invitation: 招待	Suggestion: 提案

Item prompts were composed of video-clips extracted from TV shows and films. After being edited and rendered, the chosen 38 were then compiled into one video, with each clip displayed twice for testing purposes.

Clip selection occurred during construction of a framework of pedagogic materials designed for explicit instruction of the pragmalinguistic forms and sociopragmatic features made relevant by the interaction within the clip context. The larger purpose of this instrument is to serve as a reliable pre-/post-test measure of increases in learner knowledge of specific SAs, especially those seldom found in existing instructional materials — namely, *suggestions*, *requests*, *compliments*, and *offers*. Later analysis serves to highlight shortcomings of the instrument and inform further instrument design.

Instrument Design & Construction

Each clip was cropped in length to approximately 15 seconds and edited to highlight one specific SA performed within context and restricted to two (sometimes three) speakers. Length of utterances, rate of speech and an analysis of lexical frequency, range, type and token were not performed on the speech samples. A modified Conversation Analysis approach to analyzing turn-taking guided clip selection, in that a first-pair part made relevant in some fashion a second-pair part within each clip. Notice in Figure 2 (taken from the HBO series *Six Feet Under*) the presentation of the SA *offer* in a two person, two-turn exchange.

Figure 2: Sample Video Clip Rendering



Figure 2 reveals how a brief, contextually situated clip allows for a compact prompt to elicit L2 learners' pragmalinguistic/sociopragmatic competence. (This figure was constructed for explanatory purposes and so all textual information was not present within the assessment.)

A multiple-choice paper-pencil testing format (using scantron answer sheets) was then constructed to correspond with the set of video prompts. Each clip item had one correct response and four distractors. One distractor for each item was not among the collection of 10 speech acts under investigation (notice in Figure 3 option #4 *Guess*), and was assumed absent from the conversations in all clips. These types of distractors were employed to lessen the likelihood of

successful guessing resulting in a “correct” answer and make relevant certain distinctions that might cause confusion (such as whether making an *invitation* or an *offer* constitutes making a *promise*).

The “*Offer*” conversation in Figure 2 corresponds to the set of options displayed in Figure 3. The three remaining distractors (*request*, *suggestion* and *invitation*) were chosen due to pragmalinguistic or contextual similarities to the correct response or due to a noticed confusion between certain SAs during past teaching of these pragmatics materials in classroom situations. For example, Japanese learners of English often display confusion over the difference between *offer*, *request*, and *suggestion*.

Figure 3: Item Sample Selection

	Clip 28 1. Request 2. Offer 3. Suggestion 4. Guess 5. Invitation	
--	---	--

Procedures

The instrument was administered in two contexts and to four grade levels: 2nd and 3rd year high school classes, and 1st and 2nd year university classes, all of which were intact groups streamed by their respective institutions; the sample, therefore, is not randomized.

A scantron sheet accompanied the multiple-choice answer selection, which had the following directions:

Figure 4: Participant Directions

Name _____	TOEFL _____
<p>What is happening in these conversations? Look at the selection of 15 choices below. You will have 3 minutes before the start of the videos to study the 15 words. If you wish, please use your dictionary to help you during this three-minute period. You may not speak to others at this time. While you are watching the videos, please choose the number beside the correct answer and then mark the sheet you have been given. Please mark only one answer for each question. You will be shown the clips two times.</p>	

Following the three-minute prep-time and participant questions the video was started. Twenty minutes later, the video ended.

Data Analysis

As the concept of pragmatic knowledge is not understood in uncontested, concrete terms — even to the point of how certain SAs are identified by various native speakers (NS) — it was considered crucial that a partial credit point structure be employed as part of the Rasch analysis. Therefore, 17 native speakers (13 Americans, 3 Canadians, and 1 British) sat the same instrument as the Japanese participants. Table 1 shows the NS results on the 38 items. The far-left column displays the participant numbers; the three rows at the top display the item number, the SA code, and the response coded “correct” by the researcher; and at the bottom is the percent agreement for each item.

Table 1: Native-speaker Agreement Chart

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	
	A	GG	Int	P	R	Inv	C	G	S	O	A	Inv	C	Int	GG	S	F	G	R	O	A	F	C	P	Int	G	C	S	O	G	Inv	R	S	Inv	A	R	Int		
7 0001	2	4	2	3	2	1	5	5	5	4	3	1	1	5	4	4	1	4	2	5	2	2	1	2	2	4	2	2	2	4	*	2	2	5	5	1	5	1	2
7 0002	2	4	2	3	2	1	5	5	5	4	3	4	3	5	4	4	1	4	5	5	5	2	1	2	2	4	2	2	1	*	1	1	5	5	1	5	1	2	
7 0003	2	4	2	3	2	1	5	5	5	4	5	4	3	5	4	4	1	4	5	5	5	2	1	2	2	4	2	2	1	3	2	2	5	5	1	5	1	2	
7 0004	2	4	2	3	2	1	5	5	5	4	3	1	3	5	4	4	1	4	2	5	5	5	2	1	2	2	4	2	2	1	3	2	2	5	5	1	5	1	2
7 0005	2	4	2	3	2	1	5	5	5	4	3	1	3	5	4	4	1	4	2	5	5	2	2	1	2	2	4	2	2	1	4	2	2	5	5	1	5	1	2
7 0006	2	4	2	3	2	1	5	5	5	4	*	1	3	*	4	4	1	4	2	5	5	2	1	2	2	5	2	2	4	3	2	2	5	5	1	5	1	2	
7 0007	2	4	2	3	2	1	5	5	5	4	5	1	3	5	4	4	1	4	2	5	5	2	1	4	2	2	2	2	4	3	2	2	5	5	1	5	1	2	
7 0008	2	4	2	3	2	1	5	5	5	4	5	1	3	2	4	4	1	4	2	5	5	2	1	2	2	3	2	2	4	3	2	2	5	5	1	5	1	2	
7 0009	2	4	2	3	2	1	5	5	5	4	3	1	3	5	4	4	2	4	2	5	5	2	1	4	2	3	2	2	5	3	2	2	5	5	2	1	5	1	2
7 0010	2	4	2	*	1	5	5	5	4	5	5	3	5	5	4	4	1	4	2	5	5	2	1	4	2	3	2	2	*	*	2	2	5	5	1	5	*	2	
7 0011	2	4	2	3	2	1	5	5	5	4	3	*	3	5	4	4	1	4	5	5	2	1	2	2	4	2	2	4	*	2	2	5	5	1	2	1	2		
7 0012	2	4	2	3	2	1	5	5	5	4	5	4	3	5	4	4	5	1	4	5	5	2	1	4	2	2	4	2	2	1	3	2	2	5	5	2	5	1	2
7 0013	2	4	2	3	2	1	5	5	5	4	5	4	3	5	4	4	1	4	2	5	5	3	1	2	2	3	2	*	5	2	2	2	5	5	3	5	1	2	
7 0014	2	4	2	3	2	1	5	5	5	4	*	1	3	5	4	4	2	4	5	5	2	1	*	2	5	1	2	1	3	2	4	5	5	1	3	1	2		
7 0015	2	4	2	*	3	1	5	5	2	5	1	3	*	4	4	1	4	2	5	5	2	1	2	2	1	2	2	2	*	3	4	5	5	1	2	1	2		
7 0016	2	4	2	3	2	1	5	5	5	4	3	1	3	5	4	4	1	4	2	5	5	2	1	1	2	3	2	2	1	2	2	2	5	5	1	2	1	2	
7 0017	2	4	2	3	2	1	5	5	5	2	3	1	3	5	4	4	1	4	2	5	5	2	1	3	2	3	2	2	1	2	2	2	5	5	1	5	1	2	
0	0	0	0	3	1	1	0	0	1	3	9	6	0	2	1	1	2	0	5	1	2	1	0	13	0	10	1	1	12	10	1	4	0	3	2	5	1	0	
17	17	17	14	16	16	17	17	16	14	6	11	17	15	16	16	15	17	12	16	15	16	17	4	17	7	16	16	5	7	16	13	17	14	15	12	16	17		
100	100	100	100	82	94	94	100	100	94	82	47	63	100	88	94	94	85	100	71	94	85	100	24	100	41	94	94	29	41	94	76	100	82	86	71	94	100		

These data reveal how tenuous a single researcher-coded set of pragmatic prompts can be, raising the complex issues involved in conceptualizing construct validity of pragmatic assessments and how best to adequately inform decisions regarding what set of utterances constitutes what specific SA.

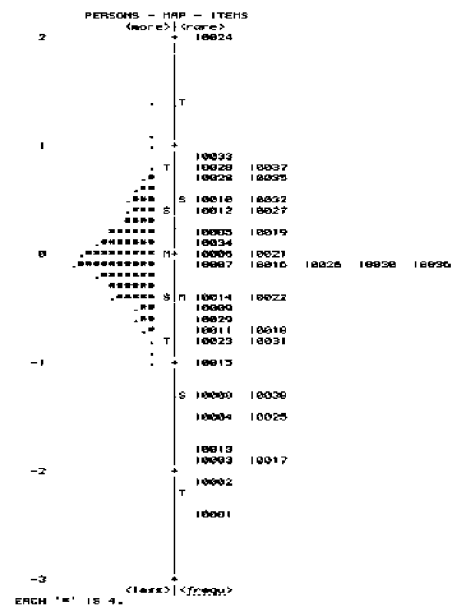
It became immediately clear item 24 should be dropped prior to item analysis, due to the fact that only 24% of the NS participants selected the designated “correct” option — with every distracter being found attractive to at least one participant. This brought the number of serviceable items to 37. Of these, only 11 were given 100% agreement, and 18 other items were agreed upon by more than 80%. Because of this, 80% (a percent deemed reasonably high for this sample of participants) was set as the arbitrary cut-off point for NS agreement constituting confirmation of the original researcher-coded set of “correct” responses. This left 8 items remaining — items that the NSs had diverged from the coded “correct” response. Partial credit

was assigned to these based on which of the “distractors” had been most attractive to the majority of NS participants. Every correct response for all 37 items resulted in a score of two points, with a partial credit point provided to responses corresponding to NS agreement.

The data was then prepped for Winsteps, along with a command file coded for partial credit (item number 24 set to be ignored), and a first run was done. The data produced a summary reporting 0.65 Person Reliability and 0.99 Item Reliability. An exhaustive iterative process of removing misfitting persons followed — based on a fit statistic cut off range from 0.70 to 1.30 with a < -2 to > 2 limit on the z-scores. The result of this process was the removal of 29 additional participants, which only raised the Person Reliability slightly. Item reliability remained at 0.99, though separations did increase.

As culling misfitting persons failed to increase person reliability estimates substantially — while, at the same time, adversely removed subjects relevant for later analysis — all participants were subsequently returned to the sample. These reliability statistics indicate that this particular sample of persons provided a much more substantial amount of information regarding the items (resulting in an extremely high item reliability) than the items offered regarding this sample population (resulting in a low person reliability). Inspecting the item map, below, provides a graphic representation of this (Figure 5).

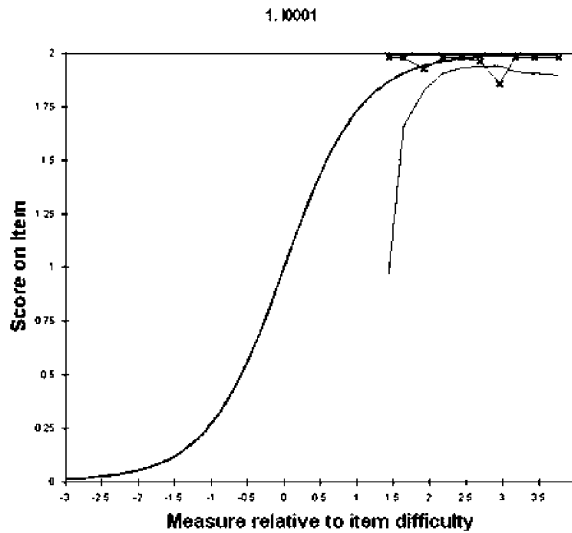
Figure 5: Item and Person Dispersion Map



The tight, bell-shaped formation of persons on the left of the scale (spreading out above and below the preset mean of zero) contrasts to the right-side item representation, with its minimal top-end spread, lower-end and middle redundancy, and empty spread at the bottom — where items fail to encounter persons of this limited ability. Effectively, items 1, 2, 3, 17, 13, 4, 25, 8, and 38 (9 of the 37 remaining items, approximately a quarter of the assessment) were too easy for the sample population. Another shortcoming revealed by the pilot run was the paucity of items at the upper range of difficulty to provide information regarding the more able persons who, apparently, “nailed” the assessment. A view of the Item Fit Order in large part confirmed this (Table 2) in that the easiest items registered as the most misfitting. However, none of the 37 items surpass 1.30 set as the threshold of outfit — with only one item, 11, registering on the cusp of overfitting model expectations.

The ICC graph for item one (Figure 6) confirms the “ceiling effect” (i.e., in that the row of Xs signifying various clusters of test-takers fall along the top, or ceiling, of the graph) noticed for items of the lower difficulty range. From previous use of various items in the classroom context, it was expected that few participants would find these of any particular difficulty. And, as such, they were placed at the beginning of the assessment to offer a few exceptionally easy items and reduce potential testing anxiety going into the test-taking experience. Item one portrays a “drink accident” at a bar in which a woman accidentally spills wine on a man’s suit. The apology this prompts is repeated and through the context alone was assumed simple and straightforward. Likewise, item two has a gift-giving prompt, where the visual cues offered (along with the “happy birthday” employed as a present is handed) might have been sufficient enough to elicit the correct response. Likewise, the NS participants showed 100% agreement on these items.

Figure 6: Apology ICC Graph



The real interest in this figure is not the across the board ceiling effects visible but that hiccups of sorts are revealed — in that for some reason persons with higher predicted levels of ability are underperforming (indicated by the pointed valley) compared to persons with lower ability levels. Especially unusual is the pronounced dip that drops below the set range for the model. This may be due to some reason other than simply a lack of knowledge — perhaps test anxiety, potentially being unprepared for the start of the administration, or possibly that a few more able participants viewed this item as being so “simple” that certain other distractors presented themselves as being more attractive — as if test-takers viewed the item as “tricky” by design.

Upon checking the list of distractor frequencies (Table 2) it becomes clear that these irregularities in the graph represent insignificant numbers of participants with predicted higher abilities who simply got it wrong. In fact, the frequencies reveal that only item one violates the progressive hierarchic assumptions predicted by the Rasch modeling of the data. Item two appears to behave as expected.

Table 2: Item Statistics for Questions One and Two

ITEM CATEGORY/OPTION/DISTRACTOR FREQUENCIES: ENTRY ORDER										
ENTRY NUMBER	DATA CODE	SCORE VALUE	DATA COUNT	%	AVERAGE MEASURE	S.E. MEAN	OUTF MNSQ	PTME CORR.	ITEM	
1	1	0	4	1	.18	.23	2.1	.06	10001	1
	4	0	1	0	.22		1.9	.03		4
	2	2	303	98	.00*	.02	1.0	-.06		2
	MISSING	***	2	1#	.02	.26		.00		
2	1	0	1	0	-.23		.9	-.03	10002	1
	3	0	7	2	-.19	.12	1.1	-.08		3
	2	0	1	0	-.19		.9	-.03		2
	4	2	300	97	.01	.02	1.0	.09		4
	MISSING	***	1	0#	-.24			-.04		

Fortunately, this non-conforming pattern of the actual data failing to meet the expected hierarchical expression of the person ability/item difficulty is found in only 6 of the 37 items (items 1, 3, 9, 12, 19, and 36 — with items 1 and 3 having been flagged earlier as part of the general ease in difficulty of the lower range).

Criterion Analysis

A more detailed discussion of which SA is represented by which item now seems called for. The collection of nine easy items at the lower end of the item difficulty graph (1, 2, 3, 17, 13, 4, 25, 8, and 38) is composed of the following set of SAs: *apologies* (1), *gift-giving* (2 and 15 — the only speech act represented by only two items, both of which contained the utterance “happy birthday,” a not uncommon English phrase for the target population), *introductions* (items 3, 25, and 38), *farewells* (item 4), *greetings* (item 8), and *compliments* (item 13). Now, if we compare these with the 9 items that represent the upper level of difficulty we start to see something of a discernable understanding of what might constitute the range of difficulties not only for the items themselves but also for SAs as a representation of distinct examples. The items at the top are *requests* (items 33 and 37), *offers* (items 10, 20, and 28), *invitations* (items 32 and 35), *suggestion* (item 12), and *compliments* (item 27).

A few things immediately interesting about this separation, and quite intuitively understandable, are that, firstly, the group of easy items that fail to find participants of a lower-level range of pragmatic knowledge compose the standard set of SAs regularly covered by the run-of-the-mill conversation textbooks — even those whose main focus is on inserting grammatical items into contrived conversational exemplars. Secondly, the higher difficulty items are composed of,

for the most part, SAs employed within more interactional types of discourse in which a certain amount of initial small talk (the kind that represents items like *greeting* someone, *introducing* yourself, and maybe giving them a “you look nice today” *compliment*). The SAs *requests*, *offers*, *invitations*, and *suggestions* are such that only a knowledge of the sociopragmatic norms of the L2 culture — in addition to the pragmalinguistic forms — would allow comprehension of (not to mention appropriate use of). These constitute a range of SAs less commonly taught and therefore in need of a more developed teaching methodology. Arguably, at some point, people enter into relationships where they have to get beyond exchanging greetings and compliments.

And, thirdly, a SA such as compliments might in fact be something of a more complex type, in that as an interactional resource it may be employed by a variety of speakers to serve distinct purposes rather than for a single set function. For instance, *compliment* item 27 shows one man saying “Nice tan” to another, who then responds with “Yeah, right.” Structurally, it is a *compliment* in the most frequently observed manner — adjective + noun, with the adjective here a part of an understood set of compliment adjectives such as “nice,” “good” and “great,” and the noun of the sort of category which is often commented upon — skin quality (as opposed to the beauty of another person’s knuckles or kneecaps). However, beyond that, it is clear (especially with the visual cues from both speakers’ facial expressions) that this particular *compliment* carries little emotional friendliness. A NS would immediately recognize that the respondent did not receive the *compliment* as one given in earnest. The character’s use of “Yeah, right” portends relationship issues beyond the literal token agreement this response could represent. For the NS participants, none of the other distractors proved more attractive, and they showed an agreement of 94% (the other 6% — i.e., 1 person — selected the distractor *greeting* here, which could be argued as relatively appropriate). As such, this *compliment*, and others like it, might be better categorized as a sub-set of *compliments* serving the purpose of criticism or sarcasm.

Motivation Survey

Based on a review of literature and careful survey of language motivation-related questionnaires previously in use, Beglar and Ono (2009) developed a range of six to nine items designed to measure each of the eight constructs detailed below. Each item was formed as a statement intended to reflect one of the constructs. For example, one of the self-confidence items read, “Compared to other students, I think I can speak English well.” Participants responded according to a six-point Likert scale of agreement, where “1” was strongly disagree and “6” was

strongly agree. The questionnaire was translated into Japanese, piloted by several native Japanese speakers, and revised accordingly based on feedback. Eight affective constructs of interest pertained to the skill of speaking in English, and included:

1. *willingness to communicate (WTC)* — the general propensity to initiate verbal communication with speakers of English when free to do so

2. *self-confidence (SC)* — the degree to which an individual has a feeling of competence when speaking English

3. *integrative orientation (INT)* — an individual's desire to interact verbally with speakers of English

4. *instrumental orientation (IO)* — the degree that an individual wishes to learn to speak English for pragmatic reasons

5. *attitude toward learning to speak English (ALSE)* — positive affect experienced while communicating in English

6. *motivational intensity (MI)* — the amount of effort that an individual is willing to expend in order to learn to speak English

7. *desire to speak English (DSE)* — the degree to which an individual has a positive attitude toward verbal communication with English speakers

8. *anxiety (ANX)* — the degree of apprehension that an individual feels when engaged in verbal communication in English (Beglar and Ono, 2008)

The motivation survey sample was initially much larger than the sample given the pragmatics test. For the purposes of this analysis, only the subjects given both tests were considered.

The response data of the 346 participants to this questionnaire were subjected to a confirmatory factor analysis using SPSS, in order to verify the number and type of components/constructs that might exist for this particular sample population.

A Principal Axis Factoring was applied, with Varimax rotation, limiting the number of factors to 8 and suppressing absolute values less than .50. In this sample ($n = 346$), the KMO Measure of Sampling Adequacy was .956. Bartlett's Test indicated a significance of $p < .05$, suggesting the relationship among the variables was strong.

As seen in Table 3, the 8 components accounted for approximately 76% of the variance.

Table 3: Total Variance Explained

Factor	% of Variance
1	35.47
2	13.52
3*	9.83
4	6.81
5	6.79
6	1.73
7	1.10
8	1.02

*Note: Factor (SC) Self-Confidence

Communalities in some items appeared lower, suggesting possible unique or random variance in these items, but most items had communalities above 0.50.

The third factor of the eight had strong loadings on questions that had been a *priori* labeled “Self-Confidence” (SC) questions. Of the nine items assumed to measure this construct, six loaded on factor three, all above .500, as can be seen in Table 4. These items did not cross-load on any other factor.

Table 4: Item loadings on Factor 3

Item	Factor loading	Communality
SC#1	.632	.53
SC#2	.627	.73
SC#4	.638	.69
SC#5	.637	.64
SC#6	.736	.71
SC#7	.761	.80

Cronbach’s Alpha measure of reliability was conducted on the six items loading on the SC factor. The results were a reassuringly high measure of .899.

Included in the motivation survey were biographical questions. Responses included a dichotomous yes/no question as to whether students had studied abroad previously (and for what amount of time), TOEFL and TOEIC scores, and, for 2nd year high school students their level of attainment on the EIKEN. For the motivation survey, students were given as much time as necessary to complete all questions, on average 10 – 15 minutes.

The results of the analysis were not expected to be generalizable beyond the sample population, as intact, non-randomized subjects were examined.

Procedure for Linear Regression Analysis on Criterion and Predictor Variables

Person logits of ability for each student were generated via Winsteps using the data from the Pragmatics Test scores. These logits were entered as a separate variable column in SPSS as indicators of person ability (i.e., receptive meta-pragmatic knowledge).

All factors on the motivation survey were converted into factor scores using the “Save as Variables” function in SPSS. The Self-Confidence (SC) factor was used in this analysis as an indication of student scores on that factor.

A linear regression was conducted on three independent (predictor) variables: academic context (i.e., what year of school the student was in), whether or not the student had lived abroad, TOEFL score (using the mean score to replace missingness), and the factor score for the Self-Confidence measure.

The rationale for this hierarchy was that students who had lived abroad in English-speaking countries were assumed to have a higher English pragmatic ability due to exposure to the language in natural settings, that students who had progressed further in school would have had more exposure to English input within the educational context and would therefore have a higher ability, and finally that a higher level of self-confidence would increase students’ pragmatic skills. Upon initial investigation, gender was found insignificant and therefore removed from analysis. A limited number of participants who reported having studied in an English-medium institution or who provided TOEIC and EIKEN scores were also removed from analysis.

Results: University students

A Pearson correlation addressed the relationship of the Pragmatics logit (criterion/ dependent variable) to the set of independent/predictor variables. The findings revealed a set of weak, though significant correlations ($p < .05$), as seen in Table 5. This indicated variation in academic context, whether a student had lived abroad, TOEFL score, and measured self-confidence were significantly correlated with the measure of the students’ pragmatic ability.

Table 5: Correlations of Independent Variables to Pragmatics Logit score

variable	r	Sig.
academic context	.328	.000*
lived abroad	.185	.000*
TOEFL	.130	.008*
SC (factor score)	.142	.004*

Note: N = 346, *p<.05

The ANOVA was also significant, $F(3, 342) = 17.18$, $p < .05$, confirming a significant portion of the variability of the Pragmatics logit score was explained by variability in the independent variables.

The regression revealed all predictors except the Self-Confidence measure to be statistically significant predictors of performance on the pragmatics measure. In line with the hypothesis, academic context had the greatest predictive power ($\beta = .29$, $p < .001$), followed by whether the student had lived abroad ($\beta = .13$, $p < .05$) and then TOEFL score ($\beta = .10$, $p < .05$). Coefficients appear in Table 7 below.

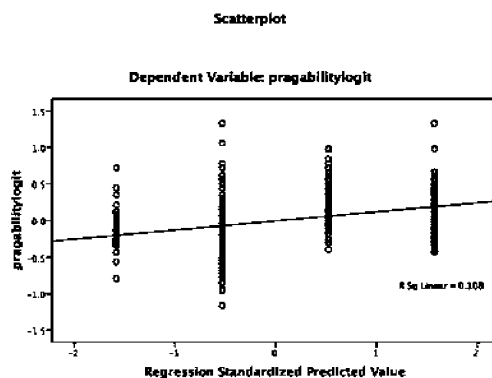
Table 6: Coefficients

	b	SE b	β
	-.93	.43	
academic context	.12	.02	.29**
lived abroad	.12	.05	.13*
TOEFL	.002	.001	.10*
SC (measure of Self-Confidence)			NS

Note: $R^2 = .13$, **p<.001, *p<.05

As indicated in Figure 8, a scatter plot of standardized residuals revealed graphically the difference in student scores based on academic context. A linear relationship was revealed, with scores statistically increasing from 2nd year high school through 2nd year university.

Figure 7: Scatter plot of standardized residuals against pragmatics



Results: High School Students

The EIKEN measure was only available for one class of students, the 3rd year group. An analysis that included the four possible predictors of EIKEN, the SC measure, and gender, on a dependent variable of the Pragmatic test score, revealed a non-significant relationship across all variables.

Discussion

The research reported was of an exploratory nature and involved the creation and pilot-test of a pragmatics instrument using both a native-speaker population as check and a relatively substantial ($n=346$) population of language learners within the Japanese high school and university education contexts. As the dependent (criterion) variable within the regression analysis, the pragmatics instrument was analyzed against a set of potential independent (predictor) variables that included gender, educational context (high school years 2 and 3, and university years 1 and 2), EIKEN score, TOEFL score, TOIEC score, residence abroad, and a 64-item Willingness to Communicate survey of motivation designed to register 8 constructs. The rational of the regression was to uncover which of these predictors might account for a learner's communicative competence (viewed as receptive knowledge of pragmatics) and its formation within the relatively monolingual (i.e., limited contact with native speakers of English outside of particular settings) educational contexts of Japanese high school and university.

This first run of a pragmatics instrument as represented by the following SAs — *suggestions*,

apologies, greetings, gift-givings, invitations, introductions, compliments, requests, offers, and leave-takings — resulted in distinctive matching of this particular set of items with this sample population. Rasch analysis is extremely sensitive to this relationship fit, and, consequently, any discussion of results cannot be generalized beyond the particular sample population.

Clearly, the instrument proved too simple for the sample drawn, which adds a positive view on a picture often presented as bleak — that proficiency abilities of Japanese learners of English remain limited. Granted, this could simply be an artifact of the sample population. Many of the high school participants aspire to attend university, and to do so within Japan for the most part requires at least a minimal display of competence in English. That such a substantial extrinsic motivating factor exists could account for participant performance on the criterion. As well, at the university level, the participants comprise a sub-set of the student population and, based on the streaming system, ranks high on language proficiency and motivation. The number of contact hours and potential opportunities to study abroad within specially designed programs far exceeds those typical of Japanese universities that are not *gaikokugo daigaku* in design. More qualitative analysis of particular sample populations would provide a fuller picture of language learning behaviors and any individual difference that could account for their performance.

The items modeled against the data as being more difficult comprise a set of vastly ignored SAs (e.g., *suggestions, offers, requests*) that could form the focus of future instructional materials. The concept of using film within the classroom is not a new one, nor are functional or task-based approaches to curricular designs and classroom activities aimed at improving students' situational use of English conversation skills. The communicative teaching approach has been around for a while. The pedagogical influence of these approaches could easily account for why the sub-group of items proved too easy for this sample of participants (though, other test-takers might find these same items more difficult, depending on their listening comprehension skills or their language experiences beyond more traditional teaching methods such as grammar translation). However, the sample's pragmatic competence evinced shortcomings on a discrete range of SAs, and this finding is most interesting in that for the first time there may be indications that pragmatic competence is more graded than previously understood. These results lean towards the view that pragmatic competence might best be viewed as a range of routines (i.e., speech acts) including those of an easily understood (potentially easily performed as well) set (e.g., *greetings, introductions, apologies, gift-giving, leave-takings*) extending on to a more complicated set of routines less clearly conceptualized within the learners' meta-

awareness of commonly employed utterances in English as a target language. It is this more difficult set (e.g., *suggestions, requests, offers, compliments*) that would best be suited for an pedagogical method focusing on explicit instruction of the pragmalinguistic forms and sociopragmatic features.

Finally, regarding the criterion, the use of a set of native-speaker participants greatly enhanced the assessment's sensitivity to the various functions this particular set of pragmalinguistic forms represent within interaction. Future research into pragmatic assessments that ignore this, more emic, perspective of pragmatics research will inevitably fail to represent the fullest possible picture of language in use.

Results on the linear regression showed that several of the predictors did prove significantly correlated with performance on the pragmatic criterion — those of individuals who had lived abroad, who had performed well on the TOEFL, and their year of education. Predictors such as attending an English-medium school, scores on the EIKEN and TOEIC, gender and any of the 8 constructs designed into the 64-item Willingness to Communicate Motivation survey failed to show significant correlation. Though the reasons for these non-significant results may prove otherwise in further studies, what should be noted is that low n-sizes for measures such as experience in an English-medium school and scores on EIKEN, TOEIC and TOEFL might very reasonably account for this. All participants, on the other hand, reported gender, and its non-significance as a predictor can be assumed a more dependable statistic in regards to representing the sample population.

The selection and data collection procedures for this set of predictor variables has provided useful insight into the choices that will drive future research, with the end goal being the testing of a structural equation model that best accounts for Japanese learners of English's communicative competence.

References

Beglar, D., and Ono, L., (2008). Motivational changes over two academic years. *JALT November Tokyo*

List of Researchers

John W. Rylander: Administered instruments, conducted Rasch analysis, and evaluated variables for S.E.M.

Phillip M. Clark: Administered instruments, analyzed data construction, and conducted Linear Regression

(ANOVA) on WTC, TOEFL, and TOIEC variables

Richard Derrah: Administered instruments, analyzed data construction, and conducted Linear Regression

(ANOVA) on EIKEN variable

J. Scott Shinall: Assisted in Rasch analysis, and was primary researcher evaluating variable power within structural equation model

The researchers would like to thank Jason Henwood, David Telega, Lisa Theisen, and Derek Weygandt, along with all other participants, for helping to make this research possible.

(John W. Rylander 外国語学部講師)

(Phillip M. Clark 外国語学部准教授)

(Richard Derrah Keimei Gakuin High School)

(J. Scott Shinall 外国語学部准教授)

