Developing a Task-based Speaking Pragmatic Assessment for Undergraduate

International Students

Shi Chen, Natalie Ferguson, & Jocelyn Rarey

Northern Arizona University

Abstract

Speaking assessments of ESL students often neglect the pragmatic aspect of speaking, or the appropriate use of specific language given the social and cultural context in which it occurs. This study was created and performed to bring more attention to the importance of teaching and testing pragmatics in speaking abilities of ESL students. Researchers created a real-world task of negotiating a group paper topic with two native English-speaking group members. Fifteen intermediate-level international ESL students at Northern Arizona University (NAU) were audio-recorded and rated by three raters on the pragmatic sub-constructs of interaction, argumentation, and language use. Ratings were scored according to the rating rubric created specifically for this study that includes four different pragmatic proficiency levels that are broken down into the three sub-constructs and their descriptions. Results of this diagnostic test showed that international students at the intensive English program of NAU are not entirely prepared by their ESL courses to perform the speech acts of suggesting or arguing through the use of pragmatic skills and awareness. With high inter-rater reliability, high sub-construct correlation, and a rating rubric that accurately depicted all of the variables we were looking to measure, this study did generally represent the pragmatic abilities of the subjects and will hopefully lead to further research.

Keywords: argumentation, ESL, group work, IEP, interaction, language use, performance-based assessment, pragmatics, rubric, task-based assessment

Developing a Task-based Speaking Pragmatic Assessment for Undergraduate

International Students

Background

The initial purpose of this study was to be able to provide Northern Arizona University's (NAU) intensive English program (IEP), the Program in Intensive English (PIE), with a test that would gauge the pragmatic competence of higher level PIE students to promote success in their studies upon matriculating into the mainstream university setting. As students enter the university, there is often a gap in their experience speaking and interacting with native speakers. As they participate in class and in group projects, this gap in their language experience becomes more evident. We sought to assess their pragmatic ability in this aspect, with a specific focus on the group project setting and argumentation. As the development of the assessment progressed, the focus shifted to the creation of the rubric. Because there is a lack of task-based performance assessment regarding pragmatic skills, the creation of a rubric to properly assess these skills was crucial to the research. We strove to create the most useful, inclusive, and concise rating rubric for the task. We found that in performance assessments, the accuracy of the rating rubric is where the actual measurement becomes indicative of the proficiency level of the test-taker for whatever skill being tested.

Research Questions

The research considers the following issues:

1. To what degree is the pragmatic task-based assessment valid from content, criterion, and construct perspective?

2. To what degree is the pragmatic task-based assessment reliable from qualitative and quantitative perspectives?

3. Do international undergraduate students (ESP PIE level 5 and 6) have the pragmatic awareness and skill to argue with native-speaking peers utilizing linguistic resources, argumentation, and interaction?

Methods

Participants

  Interlocutors. Interlocutors consisted of 2 native English speaking, female, MA-TESL students. These interlocutors were also the researchers, co-creators, and raters of the assessment. These 2 interlocutors were trained regarding the task and given loose scripts to help guide the conversation.

  Along with these native-speaking interlocutors, international non-native English speakers were recruited for the study. Participants for the assessment were level 5 and 6 PIE Engineering and Business students at NAU. Of the participants, 4 were female and 11 were male. Students' L1s backgrounds consisted of 1 Arabic, 13 Chinese, and 1 Korean. The participants had an average of 9.69 years of English Study (ranging from 8 months to 15 years). They had been residing in an ESL context (Flagstaff, Arizona) for between 3 months to 2 years. Participants were recruited on a volunteer basis only.

  Raters. The raters for the task were the three researchers and creators of the assessment. Of these, all are female, two are native English speakers and one is a native Chinese speaker with 15 years of English study. The two native English speakers are also interlocutors for the task and second-year MA-TESL students at NAU, who have completed most of the courses in the

program. One of these is currently an English 105 and English 205 instructor at NAU, and has

been for 3 semesters. The third rater is a first-year PhD student in Applied Linguistics at NAU

and holds a MAT-TESOL degree. All raters are 25 years old.

Materials

This section describes the development and design of materials created and used in the

assessment. Aside from the task, rubric, and scoring procedures, researchers also made use of a

recording device (iPhone dictation software), consent forms, and a "call for participants" script.

Design and description of task. To elicit the desired behaviors from the participants, the

task was designed to be an authentic reflection of a real-world academic setting that our

participants might encounter in the regular university setting in the future, upon leaving the PIE.

Researchers based this off of personal experience teaching composition to a combination of

international students and native-speaking students. The task was designed to mimic a

hypothetical but probable group project with two native-speaking group members and to elicit

the speech act of request making and eventually, argumentation. The task prompt was tailored to

the two different types of participants we had: engineers and business majors, in order to more

accurately test their pragmatic awareness instead of their ability to research. The topic of the

assessment fits into the curriculum test takers are studying and course instructors were consulted

in their creation. The full task is provided in the Appendix.

Design and description of rubric. Rubric design needed to be very detailed and cover

all of the sub-constructs that we wanted to measure so to assist the raters in accurately scoring

the performances analytically. Each sub-construct was given descriptions based on relevant

research.

Description of scores reported. The analytical rubric has scores 1-4 in three sub-categories or sub-constructs: interaction, argumentation, and language use. Scores from each sub-construct were combined to give each participant an overall score out of 12.

Results

Three raters gave each test taker a score from interaction, argumentation, and language use. The total score was awarded to each student by adding up the average of each sub-score. Given that descriptive statistics and frequency bar graphs can provide some evidence to evaluate the validity of the diagnostic test, the following analysis focuses on the validity evidence from statistics perspective.

Descriptive Statistics Score Report

Table 1

Descriptive Statistics for Each Sub-construct on Rating

Statistics	Interaction	Argumentation	Language Use	Total Score
K	1.00	1.00	1.00	1.00
N	15.00	15.00	15.00	15.00
Mean	2.36	2.78	2.51	7.67
Mode	2.00	3.00	2.00	8.00
Median	2.00	3.00	2.00	8.00
Min	1.00	1.00	1.00	3.00
Max	4.00	4.00	4.00	12.00
Midpoint	2.50	2.50	2.50	7.50
Range	3.00	3.00	3.00	9.00
Variance (N<30)	1.10	0.87	0.70	7.28
SD (N<30)	1.05	0.93	0.83	2.70
Skewness	0.10	-0.15	0.52	0.18
Kurtosis	-0.94	-1.12	-0.60	-0.98

As described above, the test is a task-based assessment, which mirrors a real class project setting, eliciting authentic language from students and the scoring rubric and the prompt accurately reflect the constructs that the test tries to measure. According to the correlation of the sub-constructs, each construct is closely related. Therefore, the assessment can predict international undergraduate students' pragmatic competence to some degree.

Interrater reliability is an important indicator of the reliability of the test. The interrater reliability of the total score is 0.90, which proves the test to be reliable from quantitative perspective. However, the sub-category interrater reliability of the language use is only 0.59, suggesting that the rubric needs to be changed to minimize the discrepancies amongst raters. From a qualitative perspective, the raters were quite qualified to grade test takers and were familiar with the topic and discussed the first three test takers' performances together. The interlocutors had a prepared script to help them outline possible expected interactions and were also rehearsed in order to better replicate a real discussion session with a nonnative English speaker.

Only around 20% students (a score of 10 or above) have achieved the overall pragmatic competency based on the results of the assessment. Most test takers performed well in argumentation, which means they had some content to express. However, their performance in language use and interaction are in need of improvement. Because the assessment had only 15 participants from PIE at NAU, more test takers are needed to infer a generalizable result. However, the diagnostic test is hopefully an inspiration for future research on how to improve international undergraduate students' pragmatic competence.

Relevance to PIE and Second Language learning

Most speaking tests have been focused on examining students' oral proficiency, namely the linguistic competency. Speaking assessments including the TOEFL and IELTS speaking sections, are important indicators of international students' oral proficiency achievement. However, there are not a lot of speaking tests that have been focused on assessing students' pragmatic proficiency. Pragmatics, while sometimes overlooked in speaking tests, is important to teach and test concerning international ESL students (Roever, 2009). Pragmatics are so culturally and socially bound and can lead students, while highly proficient in grammar, vocabulary, pronunciation, and fluency, to misunderstandings and even interpersonal turmoil (Clennell, 1999). Knowing in which contexts to use which language is just as important as knowing grammatical structures and vocabulary in being understood by others. This assessment is particularly important for this group of students, who are in the top levels of the PIE and are soon to matriculate into the mainstream NAU context. Engineering and business students, in particular, are expected to participate in several group projects before graduation.

References

Bachman, L., & Palmer, A. (2010). *Language assessment in practice*. Oxford: Oxford University

    Press.

Bachman, L. F. (2004). *Statistical analyses for language assessment.* Cambridge, UK:

    Cambridge University Press.

Brown, J. D. (1997). *Understanding research in second language learning*. Cambridge, UK:

    Cambridge University Press.

Clennell, C. (1999). Promoting pragmatic awareness and spoken discourse skills with EAP

    classes. *ELT Journal*, 53(2), 83-91.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.

Fulcher, G., & Davidson, F. (2007). *Language testing and assessment*. New York, NY:

    Routledge.

Kim, H.J. (2015). A qualitative analysis of rater behavior on an L2 speaking assessment.

    Language assessment Quarterly, 12, 239-261.

Liu, J. (2007). Developing a pragmatic test for Chinese EFL learners. *Language Testing*, 24,

    391-415.

Long, M. H., & Doughty, C. J. (2009). *The handbook of language teaching*. Oxford, UK: John

    Wiley & Sons.

Luoma, S. (2004). *Assessing Speaking.* Cambridge: Cambridge University Press.

Miller, M.D., Linn, R., & Gronlund, N. (2013). *Measurement and assessment in teaching* (11th

    ed.). New York, NY: Pearson.

Norris, J. M., Brown, J.D., Hudson, T., & Yoshioka, J. (1998). *Designing second language performance assessments.* Honolulu: Second Language Teaching & Curriculum Center.

Roever, C. (2009). *Teaching and Testing Pragmatics*. In Long, M. & Daughty, C. (Eds.), The Handbook of language teaching (pp. 560-577). West Sussex, United Kingdom: John Wiley & Sons.

Roever, C. (2011). Testing of second language pragmatics: Past and future. *Language Testing, 28*, 463-481.

Youn, S. J. (2015). Validity argument for assessing L2 pragmatics in interaction using mixed methods. *Language Testing*, *32*, 199-225.

Appendix

Instruments

The Speaking Prompt (Prepared for Business Students)

Participant Code:_____

You are part of a group project with your native speaker classmates and you want to bring up a change in the topic of the group project. In this project, you choose one company and research how they market their products, and you want to change the companies you'd like to search (e.g. instead of Google, you'd like to focus on Apple, Coca-Cola, Toyota, Amazon, McDonald's, H&M, Mercedes etc.)

Take some time to come up with good arguments to support your company choice. You and your partner will explain why he/she wants to keep the topic.

At the end of the conversation, you and your group members will come to some form of agreement in terms of the choosing of the company for your project. You have three minutes to explain your company choice to your group members, be sure to state and explain your reasons for your company choice.
Begin the conversation by stating your participant code (found at the top of this page), your first language, country of origin, years studying English, and amount of time spent at Northern Arizona University.

Your Notes:

The Speaking Prompt (Prepared for Engineering Students)

Participant Code: _____

You are part of a group project with your native speaker classmates and you want to bring up a change in the topic of the group project. In this project, you are going to research a renewable power plant/source, and you want to change the source/plant you'd like to research (e.g. instead of solar, you'd like to focus on hydroelectric, wind)

Take some time to come up with good arguments to support your source/plant. You and your partner will explain why he/she wants to keep the topic.

The purpose of this activity is for you both to come to some form of agreement in terms of the choosing of the source/plant for your project. You have three minutes to explain your choice to your group members, be sure to state and explain your reasons for your choice.
Begin the conversation by stating your participant code (found at the top of this page), your first language, country of origin, years studying English, and amount of time spent at Northern Arizona University.

Your Notes: