

AN ASSESSMENT USE ARGUMENT FOR SPANISH FOR THE PROFESSIONALS

C. Pardo-Ballester

Iowa State University (UNITED STATES)

This study focuses on reflective practice journal entries and compares intermediate Spanish learners' journals of their own oral performance through two different opportunities - an Oral Proficiency Interview (OPI) and a Virtual Oral Interview Classroom/based Exam (VOICES) - within the framework of a Spanish for the Professions course. Students' journals reveal that both types of assessment have their advantages in the learning process, which results in more autonomous and more efficient learners (Ziegler, 2014). Moreover, data from the two focus groups reveals information about sufficiency, relevance, beneficence, and usefulness about the Spanish speaking activities used in the Spanish for the Professions course. This study presents an assessment use argument following Bachman and Damböck (2017) that may be used to design and teach language courses using micro-formative and macro-summative oral assessments. Using an assessment use argument, claims were made about the use of VOICES and OPI. Participants in this study were 13 undergraduate students at a large public university in the Midwest of the United States and one instructor. They were enrolled in an intermediate Spanish conversation course designed for professionals. Language performance data for oral assessments from face-to-face and online formats were evaluated comparing instances of language use. The data showed that the two types of oral assessments engaged learners in different types of learning opportunities.

Keywords: assessment use argument, oral assessment, interpretive argument, reflective practice.

1 INTRODUCTION

Every semester teachers try to refine their courses in order to improve their teaching and students' learning. Sometimes teachers develop their own teaching and assessment materials and others decide to use assessments developed by third parties or they use a combination of both. As many tasks are designed for the classroom context, it is important to ensure that the validity of these tasks are comparable to authentic tasks for students to be better prepared for the workplace. At Iowa State University, students take a core course "Spanish Grammar and Conversation for Professionals" (i.e., SPAN 303B) in which instructors use a combination of assessments designed by them and by others. SPAN 303B offers intensive oral practice with the goal of improving oral proficiency. Therefore, in class, students practice speaking with materials developed by the instructor such as communicative output activities and structured output activities. Students, for this study, had the opportunity to practice speaking with authentic conversations during the American Council Teaching Foreign Languages (ACTFL) Oral Proficiency Interview (OPI) as well as with a Virtual Oral Interview Classroom/based Exam (VOICES). Moreover, self-assessments and self-reflections of their OPIs guide students toward becoming more autonomous learners (Ziegler, 2014; see Pardo-Ballester, 2019 to learn about the benefits of reflective practice). At the completion of the course, one of the learner's objectives is to be able to demonstrate speaking and listening proficiency equivalent to the ACTFL intermediate mid/high levels.

Bachman and Palmer (2010) define the assessment use argument (AUA) as "a conceptual framework for guiding the development and use of a particular language assessment, including the interpretations and uses we make on the basis of the assessment" (p.99). This study offers an AUA from Bachman and Damböck (2017) that may be used to implement classroom-based language assessments in which speaking tasks are viewed as micro-formative and macro-summative assessments. In this study, VOICES was the tool chosen as micro-formative assessment to help students with their learning process while not simply measuring its success. The purpose of using VOICES was to offer real-time feedback for students to identify their strengths and weaknesses, to empower students so they can take a greater sense of ownership over their learning by developing self-awareness, and to motivate students to practice speaking with comparable situations to those found in the OPI. The ACTFL OPI was the macro-summative assessment tool for assessing and ranking students' speaking ability with the purpose of measuring whether they reach one of the objectives of the course as well as giving students the opportunity to have an authentic conversation. Bachman and Damböck (2017) view assessment as a cyclical process in which teachers collect information about students' learning using assessments. Teachers then use the information for formative and summative decisions that lead to beneficial

consequences. These consequences may be used to improve teaching and learning. This process is done repeatedly. In this study, the speaking tasks of VOICES and OPIs were used twice in the semester. VOICES was used in the middle and end of the semester, and OPIs were conducted at the beginning and at the end of the course. Moreover, structured output activities such as gap activities and communicative output activities such as role-plays and discussions were completed weekly during the semester as micro-formative assessments. Therefore, Bachman and Damböck's (2017) cyclical process of assessment was implemented in the SPAN 303B course.

The study answers the following research questions:

1. To what extent are the speaking tasks beneficial to the students?
2. To what extent is the score-based interpretation relevant for making the appropriate decisions?

This article describes an AUA for two speaking assessments, demonstrates how backing is collected in support of the claims, and sets out ideas about how to incorporate speaking tasks to help learners improve their speaking ability. This article starts with a brief review of literature on validation theory and an assessment use argument framework for exploring the claims and backings based on (1) data collected from students' reflection journals, (2) self-assessment of their OPIs, (3) testers' scores, and (4) focus groups data.

1.1. Research on Validation Theory

From the 1990's on, validation research has been seen as the process of determining claims and constructing an argument about the inferences and uses made from test scores and evidence collection (Chapelle, Enright & Jamieson 2008; Kane 1992). Kane (1992) described the concept of interpretive argument as the three inferences (scoring, generalization, and assumptions) that are supported by relevant kinds of evidence. Later on, he extended this three-bridge formulation of the interpretive argument to a four-bridge formulation. That is, linking the target score, which can be explained by a theory-based inference to a decision about the test taker for which the score is used. With this sequence of inferences, it is possible to interpret the test-score and use the observed score to make a decision. Chapelle, et al. (2008) developed a validity argument for the TOEFL exam based on Kane's approach, but they included some changes in the interpretive argument to accommodate the TOEFL test. With their six inferences (evaluation, generalization, extrapolation, explanation, decision-making and representation), they outlined the inferences, assumptions and evidences as the base of their validity argument.

1.2. Interpretive Argument

Before discussing the structure of an interpretive argument, it would be helpful to observe that researchers have concentrated their attention on test use and the consequences of basing the validity argument (Bachman 2005, Kane 1992, Chapelle, Enright & Jamieson 2008) on Toulmin's (2003) argument approach. Bachman (2005) has proposed the AUA based on Toulmin's (2003) argument structure, which consists of making an assertion of a claim, and if this assertion is challenged, defend with data, facts, and evidence, to determine if the claim holds up. Toulmin's (2003) argument is similar to the steps a lawyer will follow to prepare a legal case. He will represent his client in a trial and argue his client's case for him. At the end of the case, the lawyer will present a closing argument and await the decision of the judge or jury. Bachman and Damböck (2017) suggest that teachers and test users need to build an argument, just as a lawyer does, to demonstrate that the intended uses of an assessment are justified. For example, if the ACTFL people say that some students in their OPI reached the intermediate-mid and intermediate-high ACTFL level at the end of a course, this is a claim. If the Spanish instructor does not accept this claim, ACTFL might provide her with OPI ranking upon which this claim was based. This claim can also be seen as an inference that is based on data or grounds as Toulmin referred to (2003). One can also ask what the connection is between students taking the OPI and reaching the intermediate-mid or intermediate-high ACTFL rank. In this example, an implied statement supports the inference "Students who finish SPAN 303B course will be ready to take the OPI and will receive an intermediate-mid rank." This statement that provides the lawfulness of the inference from data to claim is called a warrant. A warrant is the decision we take to interpret a test score. Warrants also will help support the qualities of the claims. Finding an alternative reason for students who take the OPI test such as "Not every student will reach an intermediate-mid or intermediate-high level because their grammar and vocabulary is not expanded enough and they might have difficulty understanding and responding during the oral interview" is called a rebuttal. Such a statement weakens or rejects the original claim. Then, we learn through self-reflections that "some students were not happy with the interview because the lack of vocabulary and grammar prevented having a comfortable

conversation.” This evidence supports or provides backing for the rebuttal and rejects the original claim. The argument in this example can be stated as follows:

Students reach the intermediate-mid or intermediate-high ACTFL speaking ranking at the end of the course, measured via the OPI. Since students who finish SPAN 303B course will be ready to take the OPI, they will receive an intermediate-mid rank. We can conclude that all students reach the required ACTFL level for SPAN 303B, unless some students do not have enough vocabulary and grammar to have a conversation at the level of intermediate-mid level. Some students’ self-reflections after the OPI indicated that they were not happy with the interview because of their lack in vocabulary and grammar, so we can conclude that not all students reach the intermediate-mid rank. See Fig. 1 to illustrate Toulmin’s interpretive argument within a specific context.

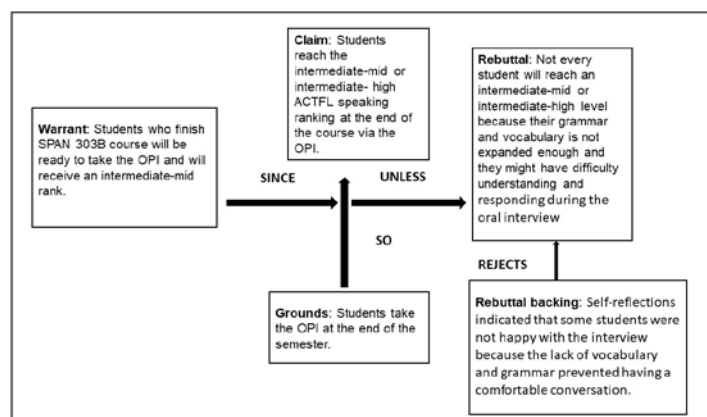


Figure 1. Toulmin’s interpretive argument for SPAN 303B argument

1.3. Assessment Use Argument

More recently, Bachman and Damböck (2017) specify an AUA to guide the process of assessment justification. An AUA is a conceptual framework for guiding the use of the OPI and VOICES speaking assessments. This framework helps us to learn how scores might be used and interpreted. Following Toulmin’s argument structure as in the example presented in Fig. 1, an AUA will include data or the information on which a claim is based, claims or statements, and backing or evidence to support the claims in the AUA. See Table 1 for a summary of the data, claims and backing associated with the AUA for the SPAN 303B course. According to Bachman and Damböck (2017) each claim in an AUA has two parts: an intended outcome (consequences, decisions, interpretations, assessment records), and one or more qualities that are claimed for the intended outcomes. The qualities or warrants, as Toulmin referred to, are: beneficence, values sensitivity, equitability, relevance, sufficiency, meaningfulness, generalization, impartiality, and consistency. The AUA presented in Table 1 is an abbreviated framework for classroom teachers, but still an effective framework to justify the use of classroom-based assessments. In the results section, two claims and backing from this AUA will be discussed in order to answer the research questions.

2 METHODOLOGY

2.1 Setting for Spanish 303B

This course prepares students to succeed in careers that involve proficiency use of Spanish. SPAN 303B focuses on refining fluency through discussions and practical experiences that reflect real-world settings in the Spanish-speaking world and among the Spanish-speaking population of the United States. The principal goal of this course is preparation of intermediate/advanced students for successful communication in the Spanish-speaking professional world. The course focuses on the ability to produce sustained and lengthy discourse in both speech and writing. Course readings, discussions, instruction, and work are entirely in Spanish. Class activities focus on building professional vocabulary and providing exercise in constructing effective and meaningful statements in the target language. This course is designed to meet the needs of students who will practice in an environment in which the use of Spanish will enhance their professional success. Students have the opportunity to reflect on some oral activities and write a report when completing two OPIs (i.e. pre and post) and four tasks using VOICES in the middle and at the end of the semester. Students have the

opportunity to listen to their own OPIs and self-evaluate their Spanish performance following the ACTFL criteria. The instructor of the course will use the result of the post OPI, along with other assessments given during the course, to make summative decisions about which students will pass the course. The instructor also makes formative decisions about the performance of her students based on VOICES and other assessments such as homework or an oral presentation.

2.2 Participants

Thirteen undergraduate students were part of this study. There were 6 female and 7 male participants. They were enrolled in SPAN 303B at Iowa State University. Their native language was English and their mean age was 19.

2.3 Oral Assessments

VOICES is a free online tool for oral assessment developed by Michigan State University. The test is composed of four random situations and each situation is followed by a prompt. Students have access to a task written in English and they hear a question in Spanish. After reading about the task, students can prepare their speech for a minute. Then, they will speak for two minutes and if they are not happy with their performance, they can redo the assignment. The prompt informs students about the Spanish question they will have to hear. These tasks are to practice presentational communication. Learners understand the situation or task because it is written in English. See Fig. 2 for the VOICES interface.

The OPI, an interview of about 15-20 minutes, emphasizes interpersonal communication which encompasses four different steps: a warm-up, level checks, probes and wind-downs. During the warm-up, the tester handles a variety of topics to have some information about the speaking ability. The interview starts at the intermediate level and most questions are about the student life and interests. During the level checks, the tester expands on same topics to see if the student can perform consistently. The tester will also ask more complex questions to get information about the type of discourse (sentences, paragraphs, extended discourse) that the students have. Then, in the probes' step, the tester will explore one level above the level checks in order to know if the speaker can sustain the performance at that level or to see if there is linguistic breakdown. He will go back and forth between the level checks and probes to avoid too much time during breakdowns. The interview finishes having a conversation at the student's linguistic level.

For the information presented above, one can say that both assessments are quite different. However, both assessments might handle the same topics but are presented in different ways. For example, in VOICES, one of the intermediate tasks is to imagine that a new language teacher will request individual meetings with his students in order to learn more about them. The goal for the teacher is to design a class to target his students' interests. With that context, students hear a prompt "Tell me about yourself." Another task at an advanced level, within the context of studying, is to ask to compare the pros and cons of on-campus or off-campus living. Students will respond to the prompt "What are the benefits of living on campus versus living off-campus?" The prompts for both of these VOICES assessments are easily asked during the OPI. For example in the warm-up, the tester interacts with the student asking him about himself (family, pastimes, studies, city, age...) and during the probes' process the tester could ask to compare two cities, or living on or off campus. One of the differences is that the OPI is based on real-world communication so the interviewee is spontaneous, using impromptu utterances, but VOICES students can prepare for a minute and can redo the presentational speech if they are not satisfied with the first result.

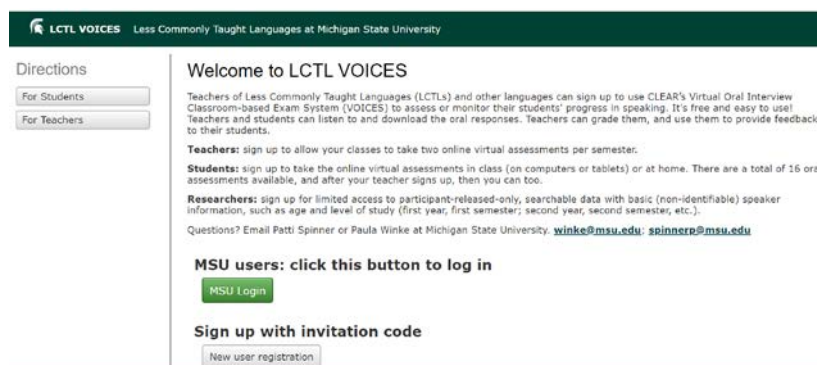


Figure 2. VOICES interface

Table 1. A Summary of the Assessment Use Argument for Spanish for the professionals.

| DATA | CLAIMS WITH OUTCOMES AND QUALITIES (WARRANTS) | BACKING |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>The ranking interpretations of the OPI based on features of 13 participants' proficiency at the beginning and at the end of the semester.</p> <p>Students who take VOICES are exposed to similar situations presented in one of the stages of the OPI (warm-up, level checks or probes.)</p> | <p>Claim 1: Intended consequences</p> <p>Outcomes: -Students will experience the assessment of the OPI and will receive feedback about their proficiency in order to know if they reached one of the objectives of the course. The scores are helpful in guiding Spanish language instruction.</p> <p>-Students using VOICES will be better prepared for the OPI at the end of the semester</p> <p>Quality: Beneficence. The consequences and decisions taken using the speaking assessments are beneficial to stakeholders.</p> <p>Claim 2: Intended decisions</p> <p>Outcome: Students' ranking based on the results of their OPI.</p> <p>Quality: Equitability. Equitability refers to the students having equivalent chances of being classified into the same group. In this study, students had the chance to take the OPI at the beginning of the semester and at the end of the semester. The tester decides to give a ranking based on the in-house OPI and therefore decides which students reach the objective at the end of the semester. The ranking of the OPI is a summative decision.</p> <p>Claim 3: Intended interpretations</p> <p>Outcomes: Aspects of students' speaking ability – accurate grammar, vocabulary, pronunciation, fluency, and communication - are relevant to the formative and summative decisions to be made. Observations of proficiency using the OPI reveal relevant knowledge in situations typical of Spanish use in interpersonal communication. The interpretations are sufficient for the formative and the summative decisions. The interpretations are meaningful with respect to the criteria of the ACTFL Proficiency Guidelines 2012.</p> <p>Qualities: Relevance, sufficiency, and meaningfulness.</p> <p>Claim 4: Intended assessment record</p> <p>Outcome: Ranking of the Pre and Post OPI.</p> <p>Quality: Consistency. The test takers' performance on the Pre and Post OPI yields essentially the same assessment records.</p> | <p>For claim 1: Data analysis of focus group asking students about their perceptions about speaking activities in the classroom</p> <p>For claim 2: Students' self-assessments for Pre and Post OPI</p> <p>For claim 3: -Students' reflections -Focus group. -Self-assessments & excerpts of discourse.</p> <p>For claim 4: Oral students' performance from VOICES yield the same assessment record as the OPI.</p> |

2.4 Reflections and Focus Groups

Students were asked to reflect on their experience after completing each OPI and VOICES assessments. Four reflections' assignments were created in Canvas. See Fig. 3 for the information they needed to provide. The reflection task was structured for successful completion of the task.

Two focus groups were conducted in the study to know about the type of activities used in the classroom and to learn about students' perceptions in relation to speaking. In one focus group, the researcher explained to students the difference between structured output activities and communicative output activities. After this explanation, she asked if students felt they were having a real conversation when completing oral tasks in the SPAN 303B course. She also asked about the context of the speaking assessments as well as the relevance, beneficence and usefulness for them. Other questions were about students' responses and if those were sufficient for the instructor to give them credit and evaluate their speaking ability. When asked about VOICES, she was interested in knowing about the difficulty

and success of the tasks, the technology (e.g., effectiveness for language practice) and if they would have achieved the same level of speaking performance without the VOICES assessments. In the second focus-group section after presenting the ACTFL proficiency guidelines 2012, students had the opportunity to self-assess their Spanish by listening to their first OPI and compare it with criteria presented in the guidelines.

When writing your reflections, think and respond the following questions using a paragraph:

How was your Spanish, based on your grammar, phonology and lexis? (Excellent, very good, good, fair, unsatisfactory: no errors, little errors, some errors, many errors, failure to understand and make myself understood)

Were you able to answer all questions, give the required information; ask the appropriate questions if required?

Was your Spanish appropriately complex, accurate, and fluent?

Was your vocabulary appropriate?

How was your experience in understanding the Spanish? Did you experience little or no difficulty in understanding, some difficulties; understanding was very limited, failing to understand adequately?

Did you use any strategies to understand the information or to answer better?

Figure 3. Instructions for self-reflections after completing the VOICES and OPIs

2.5 Data collection

Data collection included the pre-and-post OPIs with self-evaluations, the pre-and-post VOICES, self-reflections for OPI and VOICES, focus group interviews for 13 participants, and lesson plans from the instructor. Almost all data was collected during class time, except for self-reflections that were completed via Canvas outside of the classroom.

2.6 Data Analysis

The qualitative data set consisting of self-reflections and focus group interviews were compiled and entered into the software program NVivo 12 Pro. Pre-established codes were defined to have themes based on Bachman and Damböck's (2017) AUA. Codes were based on the qualities of the outcomes of the claims (e.g., beneficence, relevance...) to account for content from self-reflections. The pre-established codes for focus group interviews were based on the claims.

3 RESULTS

3.1 To what extent are the speaking tasks beneficial to the students?



Figure 4. Word clouds for beneficence from NVivo 12

Participants' responses from a focus group revealed that students consider that all speaking tasks during the course were beneficial to them. The main reason as Fig. 4 reveals was that any speaking is practice

which helps to improve their learning. They mentioned different speaking tasks, some more beneficial than others. For example, conversations were useful, but the presentational mode was the least beneficial because it was based on memory instead of impromptu conversations. Here are some perceptions from students:

“Presentations and oral assessments are probably least useful to me because I am either nervous and it throws off my speaking ability or I am just memorizing what I have to say for the presentation and not getting that real communication.” [Participant 12]

“Speaking tasks are very beneficial for me. Since I do not get to speak Spanish much outside of the classroom, my window of practice time is very slim. However, doing speaking tasks in class gives me that practice I need in order to improve my Spanish. Not only do speaking tasks help improve my speaking skills, but they also help improve reading, writing, and listening.” [Participant 8]

“I believe that the speaking tasks are extremely beneficial to me because they give an opportunity to practice Spanish in a very impromptu manner. Conversational Spanish cannot be prepared the same way essays or analysis can, and in order to improve my speaking abilities I have to practice.” [Participant 13]

Participants were also asked about their perceptions about the VOICES assessments. Responses were coded with eight themes (technology, impact, successful, improvement with VOICES, improvement without VOICES, participation, effectiveness and difficulty). Fig. 5 illustrates a word cloud with 100 frequent words with 6 as the minimum length to avoid common words such “think, but,…”.



Figure 5. Word cloud in NVivo 12 for thoughts about VOICES

“I enjoyed the tasks in VOICES. What I liked the most about voices is that I was communicating with a computer and felt that the pressure was not there as it is only a piece of technology. I was able to be more confident and able to say more of what I wanted to say. Allowing us the minute to put our thoughts together was very helpful as well.” [Participant 4]

However, two students did not agree with VOICES being beneficial to them. One was because they did not have enough exposure and the other student because the computer would not help her to get the message across. Participant 8 also agrees with participant 12 about the lack of interaction with the computer. However, participant 8 agrees that any chance to speak the language was beneficial to her:

“Because our use of it was so brief, I don’t think we got an accurate picture of its capabilities. It currently doesn’t seem particularly beneficial.” [Participant 3]

“I think this was way harder than the OPI test because there was nobody else there that knew the language that could help you out. If you didn’t know it then it was hard to even get your point across. The topics were also very unfamiliar to me and this meant less vocabulary that I could apply to the topics or questions.” [Participant 12]

“It prevents the capability of asking for help or the chance to bounce back ideas off a person. A person can tell you if they are starting to understand what you’re saying. A computer cannot do that.” [Participant 8]

3.2 To what extent is the score-based interpretation relevant for making the appropriate decisions?

The OPIs and VOICES have relevant information about students' speaking abilities. The self-reflections were structured and students expressed relevant information about grammar (lack or making mistakes), vocabulary (need of expanded lexis or being able to answer appropriately), fluency (being or not being able to maintain a regular pace or having too many pauses to look for an answer), complexity (the use of long sentences and low frequency words), pronunciation, communication (being or not being able to answer back), and comprehension (understanding the tester's questions). Self-reflections also revealed students' strategies to communicate better with the tester.

Fig. 6 illustrates students' perceptions coded in different themes for intended interpretations. These perceptions are relevant and sufficient because students reveal information about their speaking abilities and serve as back-up for the OPI final rating. Moreover, self-assessments indicated that more advanced speakers did not accurately assess their abilities after the pre-OPI as measured by the ACTFL proficiency guidelines (2012). Since only three students with lower proficiency levels out of thirteen, estimated an appropriate ACTFL speaking level, the rest of the participants tended to underestimate their speaking abilities. Regarding the post-OPI, six out of thirteen estimated their rank in the appropriate level while the rest of the participants underestimated their speaking abilities. For example, five students ranked their speaking abilities two sublevel ratings lower (i.e., four from Intermediate High to Intermediate Low and one student from Advance Low to Intermediate Mid), one student underestimated his ability only one sublevel, from Intermediate Mid to Intermediate Low. Only one student overestimated her level from Intermediate Low to Intermediate Mid. See Pardo-Ballester (2019) for descriptive statistics and t-tests for the OPIs.

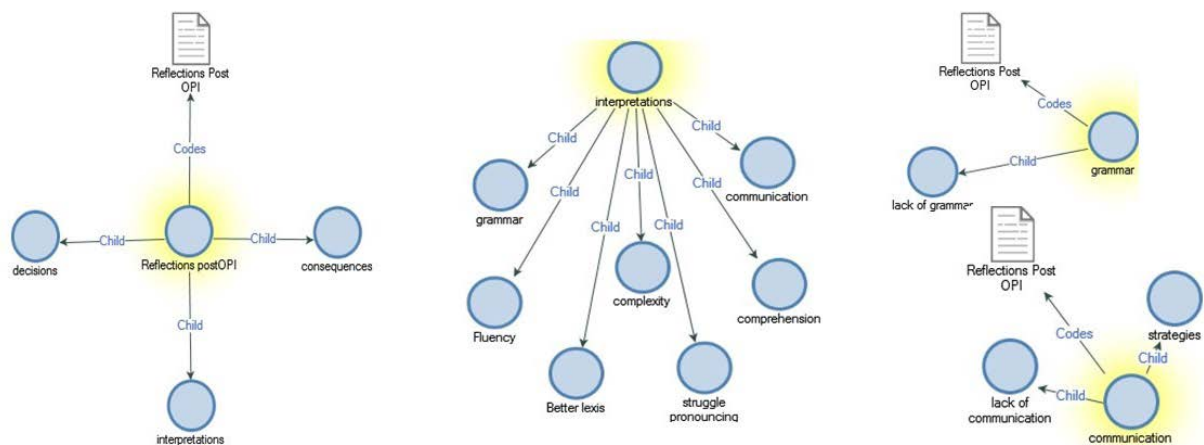


Fig. 6 NVivo Diagrams for Interpretations of Reflections post-OPI

Self-reflections empower students with the knowledge that they can improve their communication when they use effective strategies. Participants who reached the intermediate-mid or high level at the end of the semester mentioned communicative strategies such as using the context to understand the meaning, listening for key words or even ask for clarification during the conversation. Monitoring vocabulary such as self-verification is a metacognitive strategy that improves with listening comprehension success (Vandergrift & Baker, 2015):

“The strategies I used were just to rethink and ponder the prompt, as well as use context to work out the intended meaning so I could formulate my response. When trying to figure out how to say something with my limited Spanish knowledge, I would try to put the required response in other, simpler words. That strategy worked as it allowed me to convey ideas in Spanish that I was otherwise unable to find appropriate vocabulary to describe.” [Participant 3]

“While I did not understand everything perfectly or word for word, I was in the end usually able to understand what the question was or at least give an answer that would relate to the question by listening for key words in the question. I would say that this strategy helped get me through the conversation because instead of worrying about understanding everything, I only focused on finding the words I needed in order to answer the question.” [Participant 8]

"When I got stuck on a question during this OPI, I would answer in a general way to help move on to the next question or would ask for clarification on the question. Both of these methods helped lead me to a more relevant conversation and prevented awkward blips in the conversation." [Participant 2]

Chen et al. (2014) mentioned that anxiety affects learners' strategy use. This was true for participant 7 who had difficulty communicating during the interview and did not mention any strategies during the OPI except planning by studying and preparing herself for the interview. During the VOICES assessment, she used a cognitive strategy (i.e., taking notes) which relates to the assessment instead of the speaking process.

"During this OPI, I felt more prepared, but still nervous. For me, I know I get nervous when I talk to a person in another language so I tried to use a different strategy this time. I brushed over the simple words and phrases used in everyday sentences and then looked over my notes just to put some Spanish into my head before the OPI... Asking questions is a very tough subject to me, and I have always struggled with this component because questions put words in different orders and I get confused. I did use a lot of vocabulary and grammar during the question and answer part...The overall experience of an OPI is very useful and puts me in an actual conversation that could happen which I find very useful and helpful in learning a language." [Participant 7]

"...While I was talking, I know I wasn't the most fluent and most the times there were pauses or it was choppy. I am hoping to improve on the speed and fluency of my talking. My strategy this time, was to take the test first, then immediately started writing any words that came to mind. This helped me put my thoughts together to continue talking through any rough areas. Overall, I really do like the idea of the VOICES test and how it eliminates the pressure of talking to an actual person, and how I get a minute to gather my responses." [Participant 7]

Looking closely at the following excerpt of discourse during the OPI and VOICES for Participant 7 show a comparison of student's learning experience and provide backing for claim 3 for intended interpretations. The OPI interviewer gave her an intermediate role-play where she needed to go to the library to find books for a class project. Her goal was to ask the person at the front desk several questions in order to get the information she needed to use the library.

*"Hola, yo necesito un libro para mi proyecto. Y *la proyecto es sobre *la personas de Centro América y yo necesito dos o tres libros para el proyecto. ¿*De dónde *es los libros para *el Centro América? y yo necesito un libro sobre *la opuestas culturas de Centro América ¿y quién es una persona *que pregunta? ¿y cómo hago usar la tecnología para tomar un libro?"*

Literally: [Hi, I need a book for my project and the project is about the persons in Centro America and I need two or three books for my project. *Where are the books for Centro America? And I need a book about the opposite cultures of Centro America and who is the person I can ask? And how do I use technology for taking out a book?]

In the OPI excerpt, the participant tried to ask questions, but she had breakdowns with gender and number "*la proyecto*" instead of "*el proyecto*," "*la persona*" o "*la opuestas*" instead of "*las personas*" or "*las opuestas*" or using the article "*el*" when it is not necessary "*el Centro América*". She also had problems with the three questions she asked. She could not use the appropriate interrogative words, the appropriate verbs, the adequate word order or choice of vocabulary: "*¿*De dónde *es los libros para *el Centro América?*" instead of "*¿Dónde están los libros de Centro América?*" or "*¿y quién es una persona que pregunta?*" instead of "*¿y a quién tengo que preguntar?*" or "*¿y cómo hago usar la tecnología para tomar un libro?*" instead of "*¿y cómo uso la tecnología para sacar un libro?*"

Comparing the OPI excerpt with the VOICES excerpt, one can see that her knowledge asking questions was different. In VOICES she created formulaic questions and kept the questions simple in order to complete the task. She needed to ask as many questions as she could to her pretended friend who lost touch when he went abroad to study for his major. She needed "to find out what he did in the past two years, where he lives, what his plans are, and anything else." For the most part, she was successful asking simple questions, but her pace was slow and it had many pauses as she described in her reflections. Her grammar was not always accurate (problems with verbs and also constructing negative sentences) keeping it very simple. See below:

"Hola amigo ¿cómo estás? Yo estoy muy bien" [Hi, my friend, How are you? I'm very well]

"Es. Ha sido un muy largo tiempo. Yo espero que tus viajes de estudiar *están muy divertido. *¿Estás viajar a otro lado? Yo voy a Colorado muy pronto y espero que *sea España en el año twenty twenty" [Is. It has been a very long time. I hope your trips to study are very fun. Are you travelling to another place? I'm going to Colorado pretty soon and I hope to go to Spain in 2020.]

"¿Y dónde vives ahora? Yo vivo en Omaha, Nebraska, pero voy a la universidad de Iowa State" [and where do you live now? I live in Omaha, Nebraska, but I go to the university of Iowa State]

"Espero conseguir un trabajo *el año. ¿Aun vas a la escuela? La universidad ha sido muy divertido pero *es no fácil" [I hope to get a job *the year. Do you go to school? The university has been very fun, but it is not easy]

"¿Tienes novio o novia? No tengo novio y *me encanta." [Do you have a boyfriend or girlfriend? I don't have a boyfriend and I love it.]

4 CONCLUSIONS

Using the ACTFL guidelines and the OPIs to measure language proficiency was a way to measure one of the learner outcomes and support high-quality teaching. The findings provide insights into the process of preparing students to effectively analyze their speaking ability. First, the data show that students can identify their strengths and weaknesses and can set goals. Second, self-assessments used for providing a resource for students to improve their learning show that students tend to underestimate what they can do, but this tool allows participants to take responsibility for learning and understand their language proficiency. Third, evidence was gathered to argue how beneficial and relevant the OPI might be when used in a specific context. The data reported above indicate that the tests were able to separate participants into different levels of proficiency. Finally, supporting the claims for intended consequences and interpretations was briefly presented as a guide for assessment justification in the AUA.

REFERENCES

- [1] ACTFL. ACTFL proficiency guidelines 2012. Alexandria, VA: Author. Retrieved January 10, 2019, from https://www.actfl.org/sites/default/files/pdfs/public/ACTFLProficiencyGuidelines2012_FINAL.pdf
- [2] N. Ziegler. "Fostering self-regulated learning through the European Language Portfolio: An intervention mixed methods study." *The Modern Language Journal*, 98 (4), 921-936.
- [3] C. Pardo-Ballester. "The benefits of reflective practice using speaking tasks in teaching Spanish for the professionals". 12th International Conference of Education, Research and Innovation ICERI. ISBN: 978-84-09-14755-7. DOI:10.21125/iceri.2019.
- [4] L.F. Bachman and A. Palmer. *Language Assessment in Practice*. Oxford, UK: Oxford University Press, 2010.
- [5] L.F. Bachman & B. Damböck. *Language Assessment for Classroom Teachers*. Oxford: Oxford University Press. 2017.
- [6] C.A. Chapelle, M.E. Enright & J. Jamieson. *Building a Validity Argument for the Test of English as a Foreign Language*. New York: Routledge. 2008.
- [7] L. F. Bachman. "Building and supporting a case for test use." *Language Assessment Quarterly*, 2, 1-34. 2005.
- [8] S.E. Toulmin. *The uses of argument (Update ed.)* Cambridge, U.K: Cambridge University Press. 2003.
- [9] L. Vandergrift & S. Baker. "Learners variables in second language listening comprehension: An exploratory path analysis." *Language Learning*, 65 (2): 390-416. 2015.
- [10] L. Chen, R. Zhang, & C. Liu. "Listening strategy use and influential factors in Web-based computer assisted language learning." *Journal of Computer Assisted Learning*, 30, 207-219. 2014.