View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by University of Richmond

Masthead Logo

University of Richmond
UR Scholarship Repository

Latin American, Latino and Iberian Studies Faculty Publications

Latin American, Latino and Iberian Studies

11-10-2015

Increasing Language Awareness and Self-efficacy of FL Students Using Self-assessment and the ACTFL Proficiency Guidelines

Elizabeth M. Kissling
University of Richmond, ekisslin@richmond.edu

M. E. O'Donnell

Follow this and additional works at: https://scholarship.richmond.edu/lalis-faculty-publications

Part of the Latin American Languages and Societies Commons, and the Linguistics Commons

This is a pre-publication author manuscript of the final, published article.

Recommended Citation

Kissling, Elizabeth M. and O'Donnell, M. E., "Increasing Language Awareness and Self-efficacy of FL Students Using Self-assessment and the ACTFL Proficiency Guidelines" (2015). *Latin American, Latino and Iberian Studies Faculty Publications*. 151.
https://scholarship.richmond.edu/lalis-faculty-publications/151

This Post-print Article is brought to you for free and open access by the Latin American, Latino and Iberian Studies at UR Scholarship Repository. It has been accepted for inclusion in Latin American, Latino and Iberian Studies Faculty Publications by an authorized administrator of UR Scholarship Repository. For more information, please contact scholarshiprepository@richmond.edu.

Title

Increasing Language Awareness and Self-efficacy of FL Students Using Self-assessment and the ACTFL Proficiency Guidelines

Abstract

This study describes how oral language was assessed in an advanced-level college foreign language (FL) conversation course. Learners used the ACTFL Proficiency Guidelines to guide self-analyses of their oral production at intervals throughout the course. The intent was to provide opportunities for learners to develop an understanding of what constitutes oral proficiency, gauge their own progress, and set personal goals. Learners' self-analysis narratives suggested they began to notice different aspects of their speech and to better articulate their abilities and limitations. Broadly speaking, the results suggest that self-assessment of oral performance guided by the Proficiency Guidelines is an effective way to increase FL students' language awareness and self-efficacy. Pedagogical implications and limitations to this approach are discussed.

Keywords

Language awareness, self-efficacy, self-assessment, foreign language pedagogy, oral proficiency

Introduction

Most foreign language (FL) learners would argue that one of the primary objectives of language study is to improve speaking skills, that is, increase oral proficiency. But learners frequently lament their inability to communicate even after years of study and express a desire to spend more time honing oral skills (Tse, 2000). In the same regard instructors, especially those teaching conversation courses, hope learners will improve communicative ability but often find that guiding learners towards increased oral proficiency is difficult at best. Many factors contribute to these difficulties including (a) problems with the conceptualization, design, and implementation of effective courses; (b) variability in learners' initial oral proficiency, hence the need for differentiated instruction; (c) uncertainty in regard to appropriate means to measure learners' oral performance; and (d) a perceived need for greater accountability from learners for their own successes and/or failures.

Here we describe how the development of oral language proficiency was treated in one advanced-level college FL conversation course. In this course, learners used the American Council on the Teaching of Foreign Languages (ACTFL) Proficiency Guidelines to guide analyses of oral assessments made at the beginning, middle, and end of the course (http://www.actfl.org/publications/guidelines-and-manuals/actfl-proficiency-guidelines-2012). These analyses were intended as opportunities for learners to develop an understanding of what constitutes oral proficiency, gauge their own progress, and set personal goals. Through the process of self-assessment, the learners increased their language awareness and self-efficacy. Prior to describing how the course

was designed, a literature review is offered to explain the key terms, review the research on student self-assessment in language classes, and provide a rationale for choosing the ACTFL Proficiency Guidelines.

Student Self-assessment in Language Instruction

Over the last 20 years, foreign language teaching in the U.S. has become increasingly learner-centered and focused on the functional use of language (communicative competency) as opposed to mastery of lexico-grammatical structures. (See page 4 of https://www.actfl.org/sites/default/files/pdfs/21stCenturySkillsMap/p21_worldlanguagesmap.pdf). Ideally in learner-centered instruction, students become agents in their own learning process and identify what they can and cannot accomplish through the language, which enables them to set their own goals and reflect on their own progress (http://www.nclrc.org/essentials/goalsmethods/method.htm). In order to do so, in our view, learners need knowledge of the elements that constitute successful functional communicative ability. One way in which students can become better acquainted with these elements is by self-assessing their oral production. Through repeated self-assessment, learners can gradually learn to identify the many features that make up successful communication, which go beyond mere lexico-grammatical accuracy.

It is our belief that FL teaching professionals are just beginning to tap the potential benefits associated with **self-assessment**, a term that for the purpose of this study is defined as learners' ability to, "judge their own work to improve performance [by identifying] discrepancies between current and desired performance" (McMillan & Hearn 2008, p. 40). One potential benefit to well-executed practice in self-assessment is heightened **language awareness**, which in this case is defined as, "explicit knowledge about language, and conscious perception and sensitivity in language learning, language teaching and language use" (ALA, 2014).[1] Language awareness arising from the purposeful, repeated practice of self-assessment will ideally result in learners becoming more aware of their own communicative strengths and challenges, or improved **self-efficacy**, defined as "peoples' judgment of their capabilities to organize and execute courses of action required to attain designated types of performances. It is concerned not with the skills one has but with judgments of what one can do with whatever skills one possesses" (Bandura, 1986, p. 391). Stated simply, students need to know what they can and cannot do in order to strategically and intentionally improve their oral communicative performance. We believe that self-assessment can encourage development of language awareness and self-efficacy.

Research suggests that improved language awareness and self-efficacy benefit FL learners. Marsh (2008) contends that greater language awareness contributes to a deeper understanding of how language can be used to achieve specific communicative goals. There is also empirical evidence of a connection between language awareness and specific behavioral outcomes; for instance Kennedy and Trofimovich (2010) found that learners who demonstrated greater language awareness in their journals achieved more target-like pronunciation and tended to seek more contact with the target language

outside of class. In regard to self-efficacy, Bandura (1993) argues that, "there are three different levels at which perceived self-efficacy operates as an important contributor to academic development. Students' beliefs in their efficacy to regulate their own learning and to master academic activities determine their aspirations, level of motivation, and academic accomplishments" (pg. 117). Schunk and Ertmer contend that, "students with high self-efficacy are more likely to engage in activities, work harder, persist longer when they encounter difficulties, use effective learning strategies, and demonstrate higher achievement" (2000, p. 631). There is also empirical evidence that self-efficacy predicts behavioral outcomes. For instance, Mills (2007) found that self-efficacy made a unique contribution to predicting grades earned by university students (*n*=303) studying elementary and intermediate French, even when other motivational variables such as anxiety and attitudes towards French language and culture were controlled for.

Past research also suggests that self-assessment is beneficial for many reasons. In instructed learning contexts, self-assessment has been found to (a) improve student motivation and engagement, (b) lead to more meaningful and self-directed learning, and (c) empower learners to internalize the elements necessary to be successful (McMillan & Hearn, 2008; Ross, 2006). Oscarson contends that self-assessment promotes learning because it gives students training in how learning is evaluated, which is beneficial in-and-of itself. Also through self-assessment, learners are encouraged to focus on course content and their personal approaches to learning rather than simply on the results of performances (Oscarson 1989, 1997). More specifically in instructed FL learning contexts, self-assessment has been found to promote self-efficacy and a sense of control over one's learning (Baleghizadeh & Masoun, 2014; Bandura, 1977, 1984; Hsieh, 2008), resulting in higher student achievement and improved behavior, as well as raising learners' awareness of how they learn (Glover, 2011; Hsieh, 2008; Ross, 2006). In most instances, FL learners respond favorably to self-evaluation, especially when it takes the form of formative rather than summative assessment (Brantmeier, Vanderplank, & Strube, 2012; Glover, 2011). Students appear to welcome opportunities to self-assess because they like the sense of responsibility they develop toward their learning (Brantmeier et al., 2012). Another benefit of self-assessment includes the sharing of the assessment burden between the learner and the instructor, which aligns with current trends in standards-based, learner-centered instruction (McMillan & Hearn, 2008). Also, self-assessment is cost-effective and relatively easy to design, administer, and score (Brown, Dewey, & Cox, 2014; Ross, 1998).

Two meta-analyses of the empirical studies on self-assessment of various types and in various foreign and second language contexts (Blanche & Merino, 1989; Ross, 1998) have summarized the main benefits and limitations of self-assessment. Although there is much variation across empirical studies and the results are mixed, Blanche and Merino (1989) concluded that in general learners can self-assess fairly accurately, and accurate self-assessments help teachers become aware of learners' individual needs. Learners also, "find it *easier* to estimate their purely *communicative* competence level than to estimate their mastery of grammar" (p. 332). For this reason, Blanche and Merino

suggested that instructors ask students to self-assess their verbal FL performance on a regular basis through topic or theme-based activities. Their findings align with other researchers who suggest using self-assessment with tasks that are closely related to learners' potential language use scenarios, or tasks that learners are likely to encounter or can imagine themselves experiencing (Brown et al., 2014), as well as providing assessment criteria with concrete descriptions of narrowly defined linguistic situations or tasks (Oscarson, 1997). Ross's (1998) meta-analysis also found substantial, robust correlations between learners' self-ratings and criterion measures of particular language skills, namely reading, listening, and speaking. Ross then examined the effect of experiential factors on self-assessment accuracy within the context of an English FL training program for adults in Japan (*n*=254) and found that learners were more accurate when assessing functional skills with which they had experience as compared to more global measures of proficiency. Ross suggested that instructors, "design self-assessment of language learning achievement according to specific curricular content" (p. 17).

Two more recent studies are particularly relevant to our work: Glover (2011) and Brown et al. (2014). Glover (2011) introduced self-assessment in his course with first-year university ESL students training to become teachers of English in Turkey (*n*=62). The students were trained using a variety of tasks from the Common Reference Levels (CRLs), which are self-assessment tools linked to the Common European Framework of Reference for languages (CEFR) (http://www.coe.int/t/dg4/linguistic/cadre1_en.asp). Learners were asked to assess themselves and their peers' recorded oral samples using the same criteria at two points during the course. Glover concluded that the CRLs helped raise learners' awareness of both their language learning and use. He also noted a change in students' perceptions about learning and self-assessment. Their comments became longer and more specific, relevant, and critical. By the end of the term, learners were incorporating more precise terminology to describe their oral abilities, and most believed that the process of self-assessment contributed to their overall language learning. Glover argued that, "the success … seems to have been based on the extent to which students personalized their use of the statements, engaged with the tasks, and received support from their teacher and each other through training" (p. 131). The self-assessment tasks encouraged learner engagement and personal reflection, resulting in "greater self-awareness and a more realistic view of the learners' own abilities" (p. 132).

Glover's CRLs-based self-assessment tasks not only described what learners should be able to do with the target language at given levels but also contained 'can do' descriptors for each proficiency level to assist learners in determining the level most closely aligned with their productive ability. Brown et al. (2014) evaluated the effectiveness of a similar self-assessment tool developed by ACTFL, called Can-Do Statements, and measured the degree to which students' ability to self-assess their oral skills matched their oral proficiency interview (OPI) test results. Both the Can-Do Statements and the OPI are described in more detail in the following section about ACTFL resources. The students (*n*=36) in the Brown et al. (2014) study were advanced learners of Russian participating in a 12-week internship in Moscow. Their pre- and post-

OPI results were compared with their self-assessment based on the Can-Do Statements. Although their self-assessments were highly reliable, the correlation between the OPI results and the self-assessment tool was moderate. However, this study incorporated only one self-assessment in the form of 'then and now' introspection after the sojourn, and the authors suggested that students likely would have become more accurate if they had had multiple assessment opportunities with feedback.

Given the purported benefits of self-assessment, we decided to incorporate it into an advanced-level conversation course. Our approach was designed to address some gaps in the research and follow pedagogical suggestions of prior researchers. Namely, we (a) conceptualized the assessment process to be formative rather than summative, with opportunities for instructor feedback and learner reflection, more in line with Glover (2011) than Brown et al. (2014); (b) used both Can-Do Statements as well as more holistic descriptions of oral communication provided by the ACTFL Proficiency Guidelines; (c) encouraged learners to focus on communicative competence rather than grammatical accuracy as suggested by Blanche and Merino (1989); (d) linked the self-assessment prompts to specific curricular content as suggested by Ross (1998); and (e) focused on qualitative changes in learners' language awareness and self-efficacy.

Self-assessment Using the ACTFL Proficiency Guidelines

ACTFL, the national professional organization devoted to foreign language teaching in the US, was the logical place to look for self-assessment tools suited to our context of a US college Spanish FL course. ACTFL has designed several assessments, such as an online measure of multiple language sub-skills known as AAPPL (http://aappl.actfl.org/about- aappl), but the most widely recognized ACTFL oral assessment instrument and the one most relevant to an FL conversation course is the Oral Proficiency Interview (OPI) (http://www.actfl.org/professional-development/certified-proficiency-testing-program/testing-proficiency). In order to link our measure of students' oral production to specific curricular content, we created assessments that loosely followed the structure and spirit of an OPI but in modified form, as explained in the Methods section.

ACTFL-based resources that are potentially useful for learners' self-assessment include checklists of Can-Do Statements (available through NCSSFL at http://www.ncssfl.org/LinguaFolio/index.php?checklists), Performance Descriptors (http://www.actfl.org/publications/guidelines-and-manuals/actfl-performance-descriptors-language-learners), and the Proficiency Guidelines (PGs) themselves. Although the Performance Descriptors were created in 2012 as an updated and revised version of the original PGs (1984) and are intended to measure language learners' performance in response to specific instruction, we chose to provide learners with the PGs instead. Our rationale for using the PGs was that they divide language proficiency ranges into 11 sub-levels (novice low through distinguished) rather than the three broad categories (novice, intermediate, and advanced) outlined in the Performance Descriptors, thus providing greater gradation and nuance at any given level. The fact that the PGs provide narrative

for three distinct sub-levels (low, mid, and high) within each major level could help satisfy learners' need for concrete, narrow descriptions more conducive to self-assessment than broad objectives (Brown et al., 2014; Oscarson, 1997). For these reasons, we believed the PGs would potentially be more accessible and helpful to learners. We acknowledge that the PGs presumably "do not have instructional implications," (Proficiency Guidelines, p. 3) yet we found them to be appealing for the reasons we have explained heretofore.

Having designed a method of self-assessment tailored to our particular context, our research question became: What is the effect of using self-assessment informed by the ACTFL PGs on learners' language awareness and notions of self-efficacy?

Methods

Participants

Participants (*n*=13) were learners, 8 females and 5 males, enrolled in an advanced conversation course at a large, public university in the southeastern United States. An anonymous survey was sent to learners before the term began to gauge their background (e.g., years of previous Spanish study) and interests (e.g., topics and themes they wanted to explore). As previously mentioned, the survey presented Can-Do Statements to aid the instructor in the selection of level-appropriate course tasks. For instance, one intermediate-mid Can-Do Statement was, "I can discuss familiar topics in subject areas such as geography, history, music, art, science, math, language, or literature." The threshold of what would be a level-appropriate task for a learner was set as the highest level for which the learner answered more than half of these Can-Do Statements affirmatively (see Table 1). Learners also reported whether each task in the Can-Do Statements was a real-world, personal communicative goal. The instructor used learners' reported goals to inform the design of instructional tasks throughout the term.

[Table 1 near here]

Context

The conversation course was one of several options for learners pursuing advanced coursework in Spanish. This course met twice weekly for 75 minutes. The instructor, one of the authors, was a native speaker of English with near-native abilities in Spanish and twelve years experience teaching Spanish at the tertiary level. The learning objectives of the course related to oral production were that learners (a) advance in conversational speaking and (b) identify techniques for advancement through self-assessment. These learning objectives were explicitly stated in the syllabus and explained verbally to students. The course included 8 thematic units that learners chose from the pre-term survey: celebrities, transportation, sports, education, nutrition, families, travel, and prejudices. In-class activities were conversational, carried out in a variety of groupings (e.g., pairs, small groups, and whole-class) and using a variety of task types (e.g., informal conversations, interactive presentations, interviews, and debates). Each

thematic unit included multiple tasks that targeted the intermediate-low to superior ranges—appropriate levels as indicated by the pre-term survey—and presented in a low to high sequence. Thematic units also included activities in which learners were given a list of prompts identified by proficiency level. Learners were then allowed to choose the prompts that they believed would be appropriate yet challenging. Outside-of-class assignments included topical readings and videos, preparation for in-class discussion, vocabulary study, grammar exercises, personal reflections on in-class work (captured in audio or video formats), and a language journal.

Materials and Procedures

The course included three oral assessments created by the instructor to loosely follow the structure of a computer-delivered OPI. Learners completed the assessments during class time with the instructor in a language laboratory. Each assessment included 9 prompts of increasing difficulty followed by a relatively easy closing prompt. Learners responded to the prompts displayed on their computers and recorded their responses as audio files. Afterward learners performed a role-play with the instructor, which was also recorded. The recordings were submitted online through the course management system Canvas. The first assessment was conducted in week 1 and included prompts ranging from novice-high to advanced-high. The learners had not received any instruction or preparation for these initial speaking tasks, so in this way, the assessment simulated a proficiency test similar to an OPI. The second and third assessments, at weeks 7 and 14 respectively, were tailored to learners' self-rated proficiency levels so that each included prompts below, at, and above their self-reported levels. The prompts of these assessments were derived from the course themes. Thus learners had received instruction that prepared them for these speaking tasks, and in this way the assessments targeted performance rather than proficiency. English translations of example prompts are provided in Appendix A.

After each of the three oral assessments, learners analyzed their performance as homework. They were provided with the ACTFL PGs translated into Spanish (taken from Martínez Baztán, 2008, p. 154-159) and a link to the original English version online (http://www.actfl.org/publications/guidelines-and-manuals/actfl-proficiency-guidelines-2012). They were instructed to estimate the proficiency level of each prompt, describe their performance for each prompt, estimate their overall proficiency level, substantiate their claim with concrete examples from their performance and references from the PGs, list strategies they planned to implement in order to improve, and, for the last two assessments, describe progress made since previous assessments. An English translation of the instructions students received for the self-assessment is provided in Appendix B. Learners wrote their self-assessments in Spanish and submitted them electronically through Canvas.

The instructor compared the recordings with learners' analyses and provided limited written feedback via Canvas, targeting three elements. First, she noted when a student's estimation of the proficiency level of a prompt was highly inaccurate. Second,

she inserted questions like “how?” and “why?” and comments like “more” to encourage students to expand their narratives. Finally, she noted when suggested strategies were not concrete or realistic. Half of the point value for the self-assessment assignment was devoted to students’ estimation of their overall proficiency level and concrete examples to substantiate their claims (Part II, section 1 of Appendix B), hereafter referred to as their ‘self-assessment narrative.’ It is this portion of the self-assessment that is analyzed in the rest of the paper. Each self-assessment contributed 15% to the final grade calculation (45% total). Other graded elements in the course were participation (15%), language journal (20%), homework (10%), and quizzes (10%). The instructor did not evaluate or comment on the actual proficiency level of students. Thus it was students’ effort at self-assessment that in large part determined their grade, not their proficiency level per se. This grading system incentivized completing the self-assessments with care.

Analysis

The authors qualitatively analyzed learners’ self-assessment narratives using a coding scheme based on the PGs. Similar statements in the PGs’ descriptors for different proficiency sublevels were grouped into categories. The categories that emerged were comprehensibility, discourse, fluency, form, task type, topic, and sustained level (see Appendix C). Learners’ self-assessment narratives were coded according to these categories. The same comment was coded as belonging to multiple categories if appropriate. For example, a comment about not being understood because of pronunciation was coded both as a comment related to comprehensibility and form. This method of analysis captured almost all the learners’ comments, although a few did not obviously fit in any category (e.g., displeasure with hearing own recorded voice and nervousness). Learners’ narratives were compared across assessment Times 1, 2, and 3 to see changes over time.

Results

First, learners’ comments in their self-assessment narratives were sorted into ‘positive’ (of the “I am able to …” variety) or ‘negative’ (“I am not able to …”). The instances of each type of comment were tabulated to provide a quantitative overview of learners’ perceptions of their abilities. In most categories, over time learners made quantitatively more positive statements about what they could do as opposed to what they could not do (presented in alphabetical order in Table 2).

[Table 2 near here]

In addition to adopting a generally more positive orientation in each of the identified coding categories, learners’ self-assessments also evolved qualitatively in more nuanced ways. The following sections present summaries of learners’ comments related to each of the categories (in alphabetical order), focusing on the types of comments that emerged in assessment Times 2 and 3 as compared to Time 1. In the interest of space,

only the most revealing quotations from learners' narratives are used to illustrate each claim. All learners are represented in these direct quotations save two, as one turned in exceedingly brief narratives and the other failed to submit a self-assessment. Learners wrote their narratives in Spanish, but their quotations are translated into English here. The letters in parentheses after each quote are participant pseudonyms, and the numbers refer to the time of the assessment (1, 2, or 3).

Comprehensibility

In the first self-assessment (Time 1) learners tended to comment most on *whether* an interlocutor of a particular type (e.g., any native speaker) would have understood what they said. In subsequent analyses learners tended to focus more on which aspects of their speech might reduce comprehensibility, such as "I probably need to practice my pronunciation […] so that I won't be misunderstood" (KS3).

At Times 2 and 3, learners shifted from categorical descriptions of comprehensibility (or lack thereof) towards more qualified endorsements of their abilities, as in "although I can participate in a conversation with sufficient accuracy, clarity and precision, still there is confusion and errors" (JW3). Learners began to distinguish between errors that could change their intended meaning and those that would not significantly reduce comprehensibility: "I said 'I like it' instead of 'I liked it' but those are minor errors, and my idea is still clear" (KL2). They also said that breakdowns in communication are normal: "In real life, when I speak with native speakers, they understand me. Sometimes they ask me to repeat myself, but it's common that Spanish speakers aren't always going to understand advanced-low speakers on the first try" (RW2).

Discourse

The broadest category was 'discourse,' which included (a) comments about length of utterances, (b) purpose of speech (narration or description), (c) organization, (d) structure, (e) use of connectors, (f) provision of supporting facts, (g) use of time frames, (h) reliance on stock phrases, and (i) reliance on repetition of the interlocutor's utterances (see Appendix C). Throughout the semester learners noticed unnecessary repetitions in their speech, but they reported that the amount of repetitions decreased over time. Learners paid attention to the length of their utterances, and their descriptions about length became more nuanced. For instance, in her first analysis AH stated, "I spoke in paragraphs," but in her third analysis she reanalyzed her performance at Time 1 and described it differently: "I used a lot of isolated phrases that formed sentences but not complete paragraphs." Although learners commented on the variety and accuracy of verb tenses throughout the semester, their comments transitioned from simplistic accounts, such as "I used the wrong tense of the verb *to take out*" (CM1), to a more functional view of language. At Time 3 some remarked about whether their verb forms were appropriate for the task and described the communicative functions of verb tenses, as in "I showed that I could describe events in my daily life that happened in the past" (RD3).

Likewise, their notion of 'completeness' in their answers to the task prompts became more nuanced throughout the semester. At Time 1 they frequently framed their performance in the negative ("I could not…") and their analysis tended to be simplistic, commonly reporting that their responses were short or incomplete. At Time 3 learners were more positive about their ability to give 'complete' responses and, more importantly, they could explain exactly how and why their responses were 'complete.' For example, RW first described her performance as "very brief […] could have been more complete" (RW1), but at Time 3 she explains: "I spoke in detail about comparisons and contrasts […] those with advanced-low proficiency can not only give their point of view and ideas about topics in general contexts but also can employ reasoning and logic to express ideas with more details" (RW3). Throughout the semester, learners remarked on the organization of their speech as well. Again, at Time 1 their comments were mostly of the negative variety, but at Time 3 they noticed improvement in this area, as in "the transitions between topics were more fluid, although not yet at the desired level" (EG3).

In their last analyses of the semester (Time 3), learners began to make comments about elements of discourse they had not previously noticed. For the first time they noted first language structural transfer, as in "my sentence structure is from my own language sometimes instead of Spanish" (JW3); reliance on the interlocutor, as in "I have to listen, understand, and repeat what the other person has said to support my opinions" (JW3); and the appropriateness of their tone, as in "I can express emotion in Spanish and speak with the tone necessary" (KS3).

Fluency

This category included comments about fluency, flow, pauses, reformulations and self-corrections. Over the course of the semester learners reported that their fluency improved, but they also increasingly identified factors that influenced their perceived fluency, including pausing, fillers, and confidence. They often described pauses as resulting from lack of vocabulary and reported that pausing improved with time. Learners became aware of fillers in their speech and reported that they reduced their use of fillers: "My fluency has improved a lot. One of the improvements is the use of 'uh' as a filler" (CG3).

Importantly, learners' comments changed qualitatively over time to demonstrate a new understanding of fluency. For instance, one student described fluency as a characteristic that could vary across performance and tasks, as well as the effect that disfluency could have on a potential interlocutor: "During the more advanced tasks, my speech got slower. If the speaker were not a friend, he would probably give up because my explanation was so slow" (RW3). At Time 1 learners tended to characterize pauses simply as flaws, as in "it was difficult for me to answer and you can tell because there are a lot of pauses when I am thinking of what I can say" (CM1), but by Time 3 they characterized pausing with a more balanced critique, as in "pausing to think is normal, but sometimes the flow is interrupted" (GR3). At Time 1 some learners seemed to have idealistic impressions of native speakers' fluency, for instance "I think native speakers

don't use 'um' or 'eh,' rather they speak quickly and don't pause" (CG1) whereas by Time 3 they contrasted native-like and nonnative-like markers of fluency with more nuance, for instance "I think that a native speaker would move with more fluency between concepts or at least would fill these pauses with better articulated transitions" (EG3). Some learners started to report a tradeoff between fluency and precision. In addition, they began to recognize the need for compensatory strategies, as in "if I can't converse, I have phrases like 'I'm drawing a blank' to better explain my inconsistent speech" (RW3).

Form

The category of form included references to vocabulary, grammar, pronunciation, and syntax as well as general comments about accuracy, errors, and code switching. Perhaps not unexpectedly, learners made more comments about form than most other categories combined, and they tended to interpret problems in other areas as the direct result of errors in form, particularly vocabulary and verbal morphology. As with the other categories, learners reported that their form improved over time, and even though there were more negative statements ("I can not …") about form errors at Time 2 than Time 1, the proportion of negative comments to total comments in this category continually decreased (see Table 2).

Their comments about form changed qualitatively over time as well. With regards to vocabulary, learners reported that while they often lacked the vocabulary necessary to communicate what they wanted to say, their vocabulary developed over time, becoming more extensive, rich, advanced, and/or precise. Learners' focus on vocabulary also shifted from knowing *a* word or *the right* word at Time 1, as in "for me it's difficult to find the words or phrases that I need to say something well" (KS1) to finding the *best*, *most appropriate*, or *most precise* word to communicate their intended meaning. As one student eloquently stated,

> Although I can converse about many topics, it is a bit artificial, and there are many times when I lack the vocabulary to explain my points with the precision I desire […] basically, it's the nuances that are lost. I can use the language but I can't manipulate it to add a deeper meaning than what the words convey directly (EG3).

Over time learners began to identify not only the existence of errors in their speech but also the source of those errors. For example, at Time 3 several learners identified literal translations and interference from the L1 in their speech as the source of their errors, such as "in English, children need 'structure,' but I think other nouns like 'order' or 'discipline' would work better in Spanish" (RW3). Learners also began to depict circumlocution as a useful technique rather than a deficit. At Time 1 they tended to describe circumlocution as a symptom of low oral proficiency, as in "my speech has pauses, reformulations, and self-corrections when I am trying to find the right vocabulary" (JW1), but at Time 3 they described circumlocution as a communicative strategy, for example "if I didn't know a word, I used another word or phrase to express

what I wanted to say" (AH3). Learners' comments regarding self-correction followed the same trend, initially described as symptomatic of low proficiency but eventually described as a hallmark of advanced proficiency: "but there aren't many errors with verb tenses like these, and many times if I make an error, I just corrected myself immediately" (GR3).

With regards to grammar, learners made a number of comments about verb usage and conjugation at each assessment time. Although they reported minor improvement in control of verbs over time, more importantly, their analysis of the errors changed qualitatively. At Time 1 they tended to simplistically list verb errors, as in "I said 'I *was* in my car' [with the wrong verb]" (KL1). At Time 3, however, they tended to evaluate their control of verbal morphology more holistically and relate verb usage to the ability to communicate using major time frames, as in "I was paying more attention to how you can use different tenses in these situations … I was thinking and trying to use them correctly and appropriately" (GR3). Indeed, learners seemed increasingly able to disentangle form from comprehensibility and task completion.

Task type and topic

In the original PG-informed coding scheme, task type, topic, and sustained level were all coded separately. 'Task type' included references to the situations represented in the prompts, possible complications, and task variety. 'Topic' included references to type and variety of conversational topics, topic familiarity, and abstractness. 'Sustained level' captured comments about to what degree a certain proficiency level was sustained across tasks. However, as learners tended to interpret these three factors as interrelated issues, they are presented together here.

Learners reported that they completed a number of tasks successfully. At Time 1 they noted their ability to handle basic tasks and to talk about familiar topics in uncomplicated situations. Over time they reported an increase in the variety of tasks and topics they could handle well. Indeed, learners could provide many more specific examples of topics and tasks they handled well at Time 3 (32) than at Time 1 (3). The ways in which some learners categorized their abilities shifted from thematic topics, as in "I can communicate about school, interests" (RW1), towards contextualized functions, such as "I can participate in and conduct interviews" (AH3). Learners also reported that they improved their ability to handle unexpected complications, for example "one aspect of the advanced-low level is that one can maintain a conversation in unexpected situations, which I did successfully" (RD3). Over time learners became more aware of how topic familiarity and task formality impacted their performance. Some noted that they produced less language when topics were unfamiliar, as in "I can support my opinions if the topic is familiar, but I don't include many details or facts during speeches or debates" (AH3). Others noted the difficulty in shifting to a formal register, as in "I find it hard to adapt my speech depending on with whom I'm speaking, like an important adult" (CM3).

The frequency with which learners indicated the proficiency level associated with a particular topic or task increased over time, too, as did the specificity of their reasoning, for example "I had to talk with the professor about a tourist destination that I had visited and convince her to go […] this task fits with the description of the advanced-mid level. I had to describe or share information about leisure and travel" (TB3). Only five learners explicitly stated whether or not they sustained performance at a particular proficiency level (at Time 1), as in "some times I speak as if I were only at the novice level, but in general, most of the time, my level is intermediate-mid" (JW1). However, at Time 3 some learners demonstrated a nuanced understanding of the interplay between proficiency level of the task and their resulting performance. One such student commented on both discourse: "as the tasks got more difficult, my paragraph structure diminished" and fluency: "it seems to me that my speech was somewhat fluent during the first tasks, but in the advanced-mid to superior tasks, my rhythm got slower" (RW3).

Limitations

Before discussing the implications of the above-mentioned results, study limitations are described. The main drawback of this study is its size, with only 13 participants from one course. In addition, data only included self-assessment narratives; no interviews or surveys were conducted to probe student perspectives about the self-assessment process. Also, as the self-assessment was conducted in Spanish, the possibility exists that quality and/or quantity of expression might have been limited by learners' Spanish proficiency. Although all students possessed the writing proficiency to compose nuanced narratives, their Spanish proficiency still might have impacted the information each was capable of including. This potential limitation perhaps bolsters the main thesis that self-assessment is beneficial, because learners might have developed in ways that they could not articulate in Spanish. However, using English to probe further might have provided additional student insights.

Also, the present study did not concern itself with whether or not learners could accurately rate their proficiency level, as the PGs are not designed for use by untrained raters or with performance assessments. Moreover, we cannot and do not claim that learners' proficiency levels actually changed throughout the semester, as their proficiency was not directly measured. Indeed, although learners estimated that their proficiency increased by one to two sublevels during the semester, this is likely an exaggeration of their actual progress (Brown et al., 2014). We also acknowledge that learners' self-described changes could be the result not only of performing self-assessment but also potentially several other factors including instructor feedback and journaling. Lastly, the learners in this study represent a select group (i.e., advanced-level college students in one FL conversation course). Most of these learners were highly motivated to improve their oral production. For these reasons, results from this study are of limited generalizability.

Discussion

The research question asked what is the effect of using self-assessment informed by the ACTFL PGs on learners' language awareness and notions of self-efficacy. Broadly speaking, learners' narratives suggest that they began to notice different aspects of their speech and to more fully articulate their abilities and limitations throughout the course, demonstrating increased language awareness and self-efficacy. In most categories (discourse, fluency, form, task types, and topics) over time learners made quantitatively more positive statements about what they could do as opposed to what they could not do in regard to their oral performance (see Table 2). The increase in positive comments could have multiple interpretations. First, learners may have actually improved in several categories, perhaps as the result of instruction, in-class speaking practice, familiarity with the assessment procedures, and/or the exercise of self-assessment itself, but, as described in the limitations section, we make no claims to actual improvement. The increase in number of positive comments is unlikely to align perfectly with proficiency, which tends to improve quite slowly over time. Shifting from a proficiency test (at Time 1) to a performance test (at Time 2) may have boosted perceived improvement, yet the positive comments still increased in some categories from Time 2 to Time 3, both of which were performance tests. The increase in positive comments could be due to learners simply writing longer narratives at Time 2 and 3 (on average 99 more words) in response to the instructor's feedback, which potentially encouraged learners to elaborate more in their self-assessments. However, while the number of comments increased in most categories over time, the relative proportion of positive to negative comments changed as well. Thus a more likely explanation is that learners became more aware of and/or better able to articulate their abilities, which led them to focus more on what they could do rather than what they could not do. It appears that learners built self-efficacy, or their ability to judge their own capabilities in the target language (Bandura, 1986).

Qualitatively, learners' narratives became more nuanced and evidenced a richer understanding of the factors characterizing oral performance at different proficiency levels. The evolution of learners' self-assessment seemed to suggest an emerging ability to reflect on their use of language in every category, a skill noted by others as leading to learners' awareness (Hsieh, 2008; Ross, 2006). More specifically, learners began to identify specific features in their speech that impinged upon comprehensibility, fluency, and task completion. This is not to say that every student's analysis exhibited these qualitative changes in every category from Time 1 to Time 3, but each student exhibited some change in multiple categories. In the category of comprehensibility, learners' critiques became more balanced, indicating that the comprehensibility of their speech varied across tasks, and they distinguished errors that affected comprehensibility from those that did not. As for discourse, they learned to articulate what constitutes oral paragraph structure, completeness of a response, and the communicative functions of verb tenses. In addition, they began to notice features present from Time 1 yet not salient to them until Time 3, such as L1 structural transfer and tone. Regarding fluency, rather than characterize all pauses and fillers as defects, they began to recognize that even native speech is not always perfectly fluent and, therefore, they should focus more on

reducing the nonnative-like disfluencies that could negatively affect communication. As for form, learners shifted from pinpointing grammatical errors and vocabulary deficiencies to evaluating the effect of such errors on comprehensibility, describing the source of the errors, and identifying compensatory strategies. In terms of task type and topic, learners demonstrated increased awareness of how topic familiarity and formality combine to determine, in part, the proficiency level required to successfully complete a particular task and how all these factors impacted various aspects of their oral performance. All these emerging comments reflect development in various aspects of language awareness, that is "explicit knowledge about language, and conscious perception and sensitivity in language learning […] and language use" (ALA, 2014). Namely, learners developed the ability to explicitly articulate how they use language, how various language features can be described and are related, how their first language affects their second, the importance of communicative context, what constitutes proficiency, and what characterizes more advanced language.

In all the categories mentioned above, the changes over time in learners' analyses suggest that learners' perceptions of their oral abilities began to align more closely with the holistic and communicatively-oriented spirit of the PGs (awareness of what constitutes proficiency). The changes also suggest heightened self-efficacy, as learners became more aware of their specific abilities and limitations. For instance, one student commented "it's still difficult for me to talk about topics with complex information, like requirements to advance towards my degree. To get to the superior level, I would have to practice how to negotiate important situations with confidence. Right now I would not be ready to engage in something so serious" (CM3). This heightened awareness bodes well for future improvement, because learners who can set reasonable expectations and concrete goals will likely continue to improve their oral production. One student implied as much with regard to grammatical form: "although I made a ton of errors in the final, being at the advanced level does not mean that a speaker is going to speak with perfect grammar. During the upcoming semester, I'm going to continue working on accuracy in my speech" (RW3).

Pedagogical Implications

Although not without limitations, we found that self-assessments using the ACTFL PGs hold promise as a means for FL learners to increase their language awareness and self-efficacy. These benefits vary, however, depending on the learner; here some learners gained greater insights than others. Brantmeier et al. (2012) suggest that FL learners must be at a certain level of proficiency in order to accurately self-assess their performance. There was no obvious correlation between language proficiency and increased awareness in our learners, perhaps because all our learners had met some intermediate proficiency threshold. More so than proficiency level though, students' general orientation towards learning (e.g., learning style, attitudes about FL learning) likely affects how they interpret and incorporate information gleaned from self-assessments.

Others have suggested that self-assessment be tailored to learners' interests as well as their potential future language use (Alderson, 2005; Brown et al., 2014; Oscarson, 1997). Our results support these suggestions. Basing pedagogical tasks on potential uses outside the course, focusing on functional uses of language, and allowing learners to select tasks via a pre-term survey all likely increased student engagement. Also, some assessment items allowed learners to select among different prompts, which likely contributed to greater learner autonomy and self-direction (Ross, 2006). To the extent possible, we suggest that teachers using similar self-assessments allow learners to take ownership of the process, including topic selection, assessment tasks, and assessment criteria.

Researchers have suggested that learners respond more favorably to formative than summative self-assessments (Brantmeier et al., 2012). We found learners' responses in general to be quite favorable, even for the final, and therefore more summative, assessment. We never specifically asked learners to give their impressions about the use of the PGs to inform their self-assessment, but on the course evaluations at the end of the term, no one complained about their design or implementation. Learners did comment that the course required a great deal of effort, but they did not report displeasure with how their progress was assessed.

Another pedagogical issue to consider is accuracy. Although it was never our intent to judge the accuracy of learners' assessments in terms of real proficiency, one likely drawback to this method is the potential inaccuracy of learners' perceptions. Liskin-Gasparro (1987) argued that despite making good progress in a given course, the learner might not advance in measureable proficiency over the course of a semester or year of study simply due to the way the levels are described. In contrast, all learners in this study reported that they had advanced by at least one sub-level on the ACTFL scale. Although we cannot empirically confirm that learners did not advance, since we did not seek an outside, objective measure of their proficiency, it is unlikely that the learners progressed as much as they believed. Self-assessment only works as well as the preparation learners undergo prior to beginning the process. Those unskilled in the method of self-assessment tend to overrate their abilities, particularly when judging their FL oral skills (Brown et al., 2014), as seemed to have occurred with our learners. Our students might have been more accurate if they had practiced how to assess, perhaps by using the PGs to rate anonymous recordings.

In retrospect, we would recommend modifying the way in which we presented the goals of the self-assessments and how to use the PGs to our students. Learners should understand that (a) their self-assessments might not accurately reflect their true level of language proficiency; (b) assessments targeting familiar, practiced tasks will often elicit better performance than spontaneous, unpracticed tasks (true proficiency assessments); and (c) progress through proficiency levels, particularly more advanced levels, is slow. For this reason, we suggest that students be more thoroughly briefed on the continual, incremental, and slow progress they are likely to make so as to adjust their expectations accordingly. Asking students to gauge their proficiency repeatedly and at such short

intervals (1, 7, and 14 weeks) might have had the unintended consequence that they felt compelled to exaggerate proficiency advancement. Students should be encouraged to focus on their performance gains within a given sublevel rather than expect to improve their level during one semester. By deemphasizing the task of rating proficiency, and instead emphasizing introspective task aspects, we could also avoid having learners leave the course feeling confident in their ability to rate others, a potentially harmful notion given that some learners in this course hoped to be future Spanish language teachers.

Finally, we would like to add our voices to those who support learner self-assessment as a process that is easy to design, administer, and score (Brown et al., 2014; Ross, 1998). Self-assessment leads to the sharing of the assessment burden between the learner and the instructor, which aligns with current trends in standards-based, learner-centered instruction (McMillan & Hearn, 2008). Although the results from self-assessment will vary depending on the learners, the instructor, and how the information is obtained and used, in this particular case, learners' language awareness and self-efficacy increased through its use. Releasing part of the assessment burden to our learners contributed to their greater sense of control, and at semester's end those who took the process more seriously were rewarded for their efforts and most likely—through insights gained—will go on to make greater proficiency gains as well.

Future Directions

This study represents our attempt to increase learners' involvement in course planning, implementation, and most especially assessment. Although not without shortcomings, the course benefited from this approach of including learners' voices. Whether an instructor chooses to use the PGs or some other self-evaluation aid, when learners engage in periodic self-assessment of their own global performance, they are likely to gain awareness of elements that constitute communicative competency and, therefore, take a more active role in their learning. Prior research suggests that learner-centered instruction holds a great deal of promise. Yet the idea that learners take a more direct role in their instruction and assessment is frequently resisted by some instructors teaching FL in the US and, for that matter, by some students, for a variety of reasons. Our results contribute to the growing body of literature demonstrating the advantages associated with including learners in decisions about what they are taught, how they are taught, and how they are assessed. While learner-centered pedagogies are becoming more accepted in the foreign language classrooms, learner-centered assessment deserves more attention from teachers and researchers alike.

Notes

1. We acknowledge that 'language awareness' is a multifaceted construct that has been conceptualized both more broadly and more narrowly within the fields of applied linguistics (see van Lier, 2001), second language acquisition (see Leow, 2006), and education (e.g., Bilash & Tulasiewicz, 1995).

References

Association for Language Awareness (ALA). (2014). Language awareness defined. Retrieved September 10, 2014 from http://www.lexically.net/ala/la defined.htm

Alderson, J. C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. London: Continuum Publishing.

Baleghizadeh, S., & Masoun, A. (2014). The effect of self-assessment on EFL learners' self-efficacy. *TESL Canada Journal*, *31*(1), 42-58.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191-215.

Bandura, A. (1984). Recycling misconceptions of perceived self-efficacy. *Cognitive Therapy and Research*, *8*(3), 231-255.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, *8*(2), 117-148

Bilash, O., & Tulasiewicz, W. (1995). Language awareness and its place in the Canadian curriculum. In K. A. McLeod (Ed.), *Multicultural Education: The State of the Art* (pp. 49-54). Winnipeg, Manitoba: Canadian Association of Second Language Teachers.

Blanche, P., & Merino, B. J. (1989). Self-assessment of foreign-language skills: Implications for teachers and researchers. *Language Learning*, *39*(3), 313-338.

Brantmeier, C., Vanderplank, R., & Strube, M. (2012). What about me?: Individual self-assessment by skill and level of language instruction. *System*, *40*(1), 144-160.

Brown, N., Dewey, D., & Cox, T. (2014). Assessing the validity of Can-Do statements in retrospective (then-now) self-assessment. *Foreign Language Annals*, *47*(2), 1-25.

Glover, P. (2011). Using CEFR level descriptors to raise university learners' awareness of their speaking skills. *Language Awareness*, *20*(2), 121-133.

Hsieh, P. (2008). Why are college foreign language learners' self-efficacy, attitude, and motivation so different? *International Education*, *38*(1), 76-94.

Kennedy, S., & Trofimovich, P. (2010). Language awareness and second language pronunciation: A classroom study. *Language Awareness, 19*(3), 171-185.

Leow, R. P. (2006). The role of awareness in L2 development: Theory, research, and pedagogy. *Indonesian Journal of English Language Teaching, 2*, 125-139.

Liskin-Gasparro, J. E. (1987). The ACTFL Proficiency Guidelines: An update. In A. Vladman (Ed.), *Proceedings of the Symposium on the Evaluation of Foreign Language Proficiency* (pp. 19-27). Bloomington, IN: Indiana University.

Marsh, D. (2008). Language awareness and CLIL. In N. Hornberger (Ed.), *Encyclopedia of Language and Education* (pp. 1986-1999). Springer.

Martínez Baztán, A. (2008). La evaluación oral: Una equivalencia entre las *guidelines* de ACTFL y algunas escalas del MCER. Unpublished doctoral dissertation. Universidad de Granada, Spain. Available at http://hera.ugr.es/tesisugr/17457853.pdf

McMillan, J., & Hearn, J. (2008). Student self-assessment: The key to stronger student motivation and higher achievement. *Educational Horizons, 87*(1), 40-49.

Mills, N. (2007). Self-efficacy of college intermediate French students: Relation to achievement and motivation. *Language Learning, 57*(3), 417-442.

Oscarson, M. (1989). Self-assessment of language proficiency: Rationale and applications. *Language Testing, 6*(1), 1-13.

Oscarson, M. (1997). Self-assessment of foreign and second language proficiency. *Encyclopedia of Language and Education, Vol. 7*, (pp. 175-187). Dordrecht: Kluwer Academic .

Ross, J. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research and Evaluation, 11*(10) 1-13.

Ross, S. (1998). Self-assessment in second language testing: A meta-analysis and analysis of experiential factors. *Language Testing, 15*, 1-20.

Schunk, D. H. & Ertmer, P. A. (2000). Self-regulation and academic learning: Self-efficacy enhancing interventions. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Eds.), *Handbook of Self-regulation* (pp. 631-649). San Diego, CA: Academic Press.

Tse, L. (2000). Student perceptions of foreign language study: A qualitative analysis of foreign language autobiographies. *The Modern Language Journal, 84*(1), 69-84.

van Lier, L. (2001). Language awareness. In R. Carter & D. Nunan (Eds.), *The Cambridge Guide to Teaching English to Speakers of Other Languages* (pp. 160-165). Cambridge: Cambridge University Press.

Appendix A Prompts from example oral assessment

Level	English translation
Warm-up	Do you live in a dorm, an apartment, or a house? With whom do you live? Where do you eat normally – in the cafeteria, at home, or out?
Prompt 1, Intermediate-low	Option a: You want to go to the local mall this weekend and you don't have a car, but one of your friends does. Call him or her and leave a voicemail message. Option b: Tell me why you chose to attend this university.
Prompt 2, Intermediate-mid	Option a: Summarize your dietary habits to a friend (what you eat and don't eat and why). Option b: Talk about your favorite sport. Choose a player of that sport or your favorite sports team and tell a friend how it's going for them this season.
Prompt 3, Intermediate-high	Tell a friend about a fantastic day or horrible day that you had in high school.
Prompt 4, Intermediate-high	Chose a type of diet (e.g., vegetarianism, Atkins, South Beach) and explain to a friend what it is like.
Prompt 5, Advanced-low	Talk to a friend about some current event, explaining it in detail and its relevance.
Prompt 6, Advanced-mid	Explain to a friend, step by step, how you prepare your favorite meal.
Prompt 7, Advanced-mid	Tell me about the greatest challenges and obstacles in your education.
Prompt 8, Advanced-high	Option a: Explain and defend OR refute this hypothesis to me: If the university stopped giving athletic scholarships and funding sports in general, the university would benefit. Option b: You are in a 'town hall' meeting about drunk driving in our city. You are representing students of the university. Talk about the problem and offer some solutions.
Role-play	You stayed late working in the office. No one else was there. You left to get something to eat and now that you've returned, you're locked out. You aren't carrying any form of ID. Explain the situation to the security guard and ask her to let you in.
Cool-down	What do you plan to do after graduation?

Appendix B Instructions and rubric for self-analysis (English translation)

Part I: The tasks

a. Identify the proficiency level of each task. (15 points)

b. Describe your performance on each task (10 points):
    a. Were you able to complete the task satisfactorily, expressing yourself appropriately for the communicative context and interlocutor?
    b. Identify your level of discourse:
        i. novice (isolated words)
        ii. intermediate (isolated sentences)
        iii. advanced (paragraphs)
        iv. superior (extended)
    c. Identify your level of accuracy:
        i. novice (can be difficult to understand)
        ii. intermediate (can be understood by those who are used to speaking with nonnative speakers)
        iii. advanced (can be understood by any native speaker)
        iv. superior (errors are those made by native speakers

Part II: Proficiency

1. Based on the ACTFL Proficiency Guidelines, determine your overall proficiency level. Give lots of specific examples to justify your evaluation. (50 points)

2. Compare your performance on the last oral assessment with this one. Have you improved? Provide specific examples. (15 points)

3. Enumerate various specific techniques that you plan to use to increase your current proficiency level. (10 points)

Appendix C Coding categories derived from ACTFL Proficiency Guidelines (with example of Advanced Low)

Category	Descriptors in ACTFL Proficiency Guidelines for advanced-low
Comprehensibility Keywords: misunderstandings, interlocutors' understanding	They contribute to the conversation with sufficient accuracy, clarity, and precision to convey their intended message without misrepresentation or confusion. Their speech can be understood by native speakers unaccustomed to dealing with non-natives, even though this may require some repetition or restatement.
Discourse keywords: length, type (narration, description), time frames, organization, structure, supporting facts, full accounts, linking, connection, repeat interlocutor, learned phrases	They demonstrate the ability to narrate and describe in the major time frames of past, present, and future in paragraph-length discourse with some control of aspect. In these narrations and descriptions, they combine and link sentences into connected discourse of paragraph length, although these narrations and descriptions tend to be handled separately rather than interwoven. Responses they produce are typically not longer than a single paragraph. Their dominant language may be evident in […] the oral paragraph structure of that language.
Fluency/Flow keywords: pauses, reformulations, self- corrections, stock phrases	At times their discourse may be minimal for the level, marked by an irregular flow, and containing noticeable self-correction.
Form keywords: accuracy, code switching, vocabulary (specificity, extensiveness, L1 transfer, false cognates, literal translations, re-phrasing, circumlocution), grammar (verbs, etc.), pronunciation, syntax	Their speech is typically marked by a certain grammatical roughness (e.g., inconsistent control of verb endings). Their dominant language may be evident in the use of false cognates, or literal translations. Their vocabulary often lacks specificity. Nevertheless, they are able to use communicative strategies such as rephrasing and circumlocution.
Task types / Functions keywords: situations, complications, variety, ask and answer questions, request information	They are able to handle a variety of communicative tasks. They can handle appropriately the essential linguistic challenges presented by a complication or an unexpected turn of events.
Sustained Level keywords: level of the task, deterioration, break down, uneven performance	The overall performance of the Advanced-level tasks is sustained, albeit minimally. When attempting to perform functions or handle topics associated with the Superior level, the linguistic quality and quantity of their speech will deteriorate significantly.
Topics keywords: types, variety, familiarity, abstractness	They are able to participate in most informal and some formal conversations on topics related to school, home, and leisure activities. They can also speak about some topics related to employment, current events, and matters of public and community interest.

Table 1. Results of pre-term survey

Student	Self-reported initial proficiency based on Can Do Statements	# Years high & middle school Spanish	# Spanish courses taken above intermediate	Time studied abroad in Spanish	L1(s)
1	Intermediate-low	6	3	None	English
2	Intermediate-low	4	4	None	English
3	Intermediate-low	5	10	None	English
4	Intermediate-low	5	3	None	English
5	Intermediate-low	2	1	None	English
6	Intermediate-mid	5	2	None	English
7	Intermediate-mid	4	2	None	English
8	Intermediate-mid	Not reported	3	None	English
9	Intermediate-high	4	2	1 Semester	English
10	Intermediate-high	5	4	None	English
11	Advanced-mid	Not reported	2	None	English & Spanish
12	Advanced-mid	5	4	None	English, Spanish, & Italian
13	Advanced-high	Not reported	3	6.5 Years	English

Table 2. Number of comments by time of test, descriptor category, and orientation

	Time 1			Time 2			Time 3		
	Positive	Negative	Sum	Positive	Negative	Sum	Positive	Negative	Sum
Comprehensibility	11	3	14	12	7	19	5	3	8
Discourse	4	10	22*	16	11	39*	36	9	57*
Fluency	4	10	14	12	13	25	20	6	26
Form	7	44	51	25	63	88	40	34	74
Task Type	12	3	15	12	2	14	30	7	37
Topics	14	1	15	23	1	24	55	1	56
Sustained Level	3	2	5	3	3	6	3	5	8

*Includes comments about sentence length, which were neutral