

An Interview with Lyris Wiedemann
Coordinator of the Portuguese Program
Stanford Language Center

PLJ: How do you evaluate and assess your Portuguese classes?

In order to understand the nature of our evaluation system, it is necessary to understand our context and even a bit of our history. At Stanford, language programs are neither independent nor are they branches of a corresponding literature or culture department, as it happens in most colleges. Instead, like every language taught at the university, the Portuguese language program is part of the Stanford Language Center, and thus shares with other language programs the same general guidelines for objectives, curriculum, and, naturally, assessment and evaluation.¹ Like the other programs, Portuguese is proficiency-oriented and national standards-based, and every step of the teaching-learning process is built on these cornerstones. We provide our students with foreign language capabilities that enable them to live, work, study, and research in a Lusophone country. We aim to prepare them not only for interaction at the colloquial and personal levels, but also for full engagement on issues of mutual interest, which requires more advanced linguistic abilities.

PLJ: What does this mean in concrete terms for teaching and evaluation?

This approach underlies all of our activities, both from a macro and a micro point of view. To build these competencies, we have always in mind the 5 Cs of the National Standards.² Rather than considering language learning simply a matter of acquiring linguistic abilities, our communicative approach highlights the establishment of connections and comparisons between

¹ Throughout this interview, *assessment (avaliação formativa)* is used as a process-oriented, cooperative process that informs instructors and students to what extent learning is taking place, and whether there are changes that could be introduced to optimize the learning experience; *evaluation (avaliação somativa)* is used as product-oriented, judgmental measurement, with the purpose of rating the performance of the student in relation to a previously established target. See, for instance, T. Angelo and K.P. Cross. *Classroom assessment techniques a handbook for college teachers*. (San Francisco: Jossey-Bass, 1993), 427.

² <http://www.actfl.org/i4a/pages/index.cfm?pageid=3392>

languages and cultures, as well as the discovery and gradual knowledge of the communities that speak Portuguese.

Thus, in all of our classes, programs, and daily teaching-learning, we emphasize *doing* rather than *knowing*. Our evaluation centers on this, as well. We try to make sure that students learn to speak, listen, read, and write in ways that are immediately useful in a real-world setting, and this approach is embedded in everything we do.

PLJ: How is this done, in practical terms?

We use the time in our classes almost exclusively for practice. I jokingly tell my students that I am their personal coach, not their instructor, but this assertion contains our philosophy. Every week students get a syllabus which outlines day-by-day what they should know and have prepared when they come to their daily classes. It also describes what they are expected to be able *to do* after that class. This means that we first define what the students must accomplish, what the targets are in terms of language competence, and then we organize the route to accomplish those targets as steps of a staircase. We depart from where the students are at a given stage, and find ways to map their acquisition journey so that they achieve those previously-defined targets.

PLJ: But does every individual instructor in the program have to do that? It appears like a colossal task.

No, it is not done individually, although of course instructors need to tailor content and activities to match individual classes. We have a very effective guiding light. When the Language Center was formed in the mid-nineties, one of the first actions of the newly appointed director, Elizabeth Bernhardt, was to gather a small group of instructors to draft a description of target abilities for

first- and second-year Spanish and Portuguese. This happened at the end of my first year at Stanford, and I was fortunate enough to be a member of this team. The resulting document, revised several times since then (we now have separate Spanish and Portuguese documents), establishes the guidelines for the syllabi of all of our courses, while also allowing us to maintain a bird's eye view of the program as a whole. It can be accessed by anyone at the Language Center site.³

Based on these curricular documents, we know what students are expected to develop in each level/class, and easily define how to assess both an individual student and a group at any time, as well as design final evaluations. This is a very useful and productive tool for a language coordinator. Moreover, one of the most important aspects of the Language Center is the fact that all language programs reflect the same language acquisition and teaching philosophy. Since each language program creates its own curricular documents, the pragmatics of each language and culture, as well as the relationship between genre and function inherent to the language are respected. As language systems are different, each one of the documents is unique, but they preserve the same backbone, which reflects a clear characterization of our vision of language acquisition. This includes defining what the students can do in terms of the interpersonal, interpretive, and presentational abilities at each level. It is important to remember that language competence always includes these three dimensions.

At the class level, the curricular documents simplify enormously the task of the instructor, guiding him/her as to what kinds of questions to ask in assessment tasks and defining what the student should be able to do, both on a daily basis and by the end of the program. These documents are beneficial for the students, as well. Since students know what they should be able

³ <https://www.stanford.edu/dept/lc/language/requirement/curriculum.html>.

to do by the end every class, our weekly program also guides them on a path of self-assessment and self-evaluation. For instance, we want them to start asking themselves: can I refer to the days of the week? Can I express time? Can I interrupt someone who is talking? Can I state and justify an opinion?

PLJ: And what was the basis for defining this backbone? How do you set the final goals at each level?

Our basis is research and theory on language nature, acquisition and learning, and discourse functions. As stressed by the Language Center, our general orientation is adaptive, compensatory, and developmental, not additive. So when we revise the Portuguese program curricular documents and design our general course syllabi, we are oriented by what we as language scholars know about language acquisition and learning. For instance, we know that beginners tend to repeat formulae and memorized chunks of sentences, and that the next step is going beyond that to create with language, by recombining learned elements to convey meaning. These recombinations are then interconnected at the next level of complexity, what we call speaking and writing in paragraphs. Our final evaluation does not focus on the student passing an exam, like the baccalaureate or some national language exams, but on assessing the level of language acquisition of each student in comparison to the target levels. It is truly developmental, which at first can be difficult for instructors not accustomed to defining levels of acquisition and using a developmental scale as a part of their evaluation. We do not prepare students to “pass an exam,” but rather to use the language proficiently.

PLJ: I can identify ACTFL's views on your description. What is the relationship between the Portuguese program, the Language Center and ACTFL?

We are intrinsically aligned with ACTFL in their description of language competence and acquisition, which is firmly grounded on solid research developed over many years by ACTFL and the Interagency Language Roundtable (ILR) scale.⁴ In addition to following the ACTFL proficiency guidelines⁵ as main reference, we define our class-level targets using the ACTFL scale. For example, in our accelerated classes, students should reach the Intermediate Low level or beyond by the end of the first quarter, and the Intermediate Mid or beyond by the end of the second quarter.

PLJ: What instruments do you use to verify whether students have reached these goals?

Our final exam for all Portuguese classes during the fall and winter quarters is an Oral Proficiency Interview (OPI) and a Writing Proficiency Assessment (WPA), conducted and rated by trained instructors. During the course, other evaluation instruments include vocabulary tests, communicative tests, oral comprehension tests and impromptu writing. Vocabulary tests and impromptu writing are done in class, while chapter tests are taken at home, so that we can use class time for communicative activities. The final evaluation of each course combines all of these elements. There are also numerous activities of assessment, such as, starting on the second quarter, weekly correction of a vocabulary list compiled by the student, the first correction of each composition, and so forth.

In terms of evaluation, each year at the end of the spring quarter we use an identical Simulated Oral Proficiency Interview (SOPI) and a computer-based Writing Proficiency Assessment (WPA) common to all languages, which is grounded on the ACTFL guidelines for writing.

⁴ <http://www.govtilr.org/Skills/IRL%20Scale%20History.htm>

⁵ <http://www.actfl.org/profguidelines2012>

PLJ: What makes the spring quarter different? Why do you use a different instrument in the spring?

The SOPI, as you know, is an OPI on the computer, so to speak: it uses the same structure of the OPI and also the same scale for rating. The SOPI for Portuguese was originally developed in the nineties by the Center for Applied Linguistics⁶. Since then, a few institutions like Stanford have developed their own SOPIs, keeping the same basic structure of the original instrument. We have several forms of it, and have been using them successfully for many years.

There are several advantages to using the SOPI for at least one quarter each year. First, it eliminates the interviewer variable, which makes results more reliable and more easily comparable than comparing situations with different interviewers. Since all students take the same computer-based test, it makes it easier to carry out statistical analysis. Second, it also makes it possible to compare the performance of different language programs. Of course you can also do this with the OPI, but if you want to increase the reliability, then eliminating these two intervening variables is a great idea. Another advantage is that it can be applied to an entire class at the same time. Finally, using this “impersonal” instrument allows the Director of the Language Center to fulfill her obligation of annually reporting data for all language classes to the Committee on Undergraduate Standards and Policy in a more concise manner.

PLJ: But can a computer-based example be as effective as a human interviewer when eliciting samples from speakers?

There are a good number of studies comparing the OPI and the SOPI, or other face-to-face/telephone interviews with computer-based interviews.⁷ The overwhelming majority of them

⁶ Stansfield, C.W., Kenyon, D.M., Paiva, D., Doyle, F., Ulsh, I., & Cowles, M.A. (1990). “Development and validation of the Portuguese Speaking Test.” *Hispania* 73:3 (Sep. 1990): 641-651.

⁷ See, for instance, Shohamy, E., Gordon, C., Kenyon, D.M., & Stansfield, C.W. (1989) “The development and validation of a semi-direct test for assessing oral proficiency in Hebrew” *Bulletin Of*

have found a very high correlation between the two tests. Nevertheless, I think more people still prefer a face-to-face interview. Informally I can say that most of our students do. The main issues are the time-limitation and the inflexibility of the SOPI tasks. The ideal solution is probably using adaptive computer-based tests. I will not discuss those because we are not using them yet.

PLJ: What do you learn from assessment and evaluation of your Portuguese program? How do you use that knowledge in the program?

Basically assessment tells both instructor and students how learning is proceeding and how it could be enhanced. For instance, by observing closely the daily work of each student, the instructor (or “personal trainer”) can easily assess the student by focusing on the “what you should be able to do” line in the daily program as s/he moves around the class. This can be done very simply, for instance, by filling in “satisfies”/“does not satisfy” on the rubric or—to use ACTFL terms—“meets”/“exceeds”/“does not meet.” The same thing happens with the first correction of the compositions and the weekly correction of the vocabularies. All of these allow us to understand the strengths and weaknesses of individual students and classes, and lead us to reformulations or remedial work, as necessary.

At this point I must stress again that our general syllabus is actually just a general road map to guide the student. It is changed several times during the quarter, and students easily get accustomed to the idea that what is really important is the weekly syllabus, because it takes into consideration where the class actually is and how it is progressing.

Hebrew Higher Education 4(1); Kuo, Jane, & Jiang, Xixiang (1997). “Assessing the Assessments: The OPI and the SOPI.” *Foreign Language Annals*, 30(4): 503-12; Mikhailova, Julia V. *Comparison of Interpersonal and Presentational Description in Russian Oral Proficiency Testing* (Doctoral Dissertation), Ohio State University, 2005, to mention a few.

Nevertheless, the general syllabus is very important because it defines the evaluation tasks for the final grade. From these, i.e. the tests, the impromptu writing, the final OPI/SOPI and WPA, we gather the concrete data that result in student grades and curriculum/syllabus restructuring. The evaluation data tell us how individual students and the class in general performed in relation to the target established. For instance, if a considerable number of students did not reach the target level in a class, then there is something to be improved either in the content or the tasks of that class, its management, or the program. We may decide, for instance, to change the order in which we present the materials, look for failures to be corrected in our methodology, replace teaching materials, or work directly with the instructor to overcome specific difficulties.

This information is complemented by course and instructor assessment and a final evaluation. The former is carried out around midterm and is strongly encouraged by the Language Center, but is not mandatory. The midterm assessment can be conducted in paper or electronic format by the instructor, or by a specialist from the Center for Learning and Teaching, who visits the class, meets with the students without the instructor, and then discusses the students' feedback with the instructor. The final evaluation is carried out electronically at the end of each quarter, before grades are published; it is mandatory and consistent across the entire university. Both are anonymous. The program coordinator reviews the former, and suggests corrections as appropriate. The Director of the Language Center reviews the latter, and considers it as a part of each instructor's annual evaluation.

The annual report of the Language Center to the Committee on Undergraduate Standards and Policy is also a source of evaluation for the Portuguese program, as it allows us to see how Portuguese is doing in comparison to the other languages. In addition, it is a source of diachronic

information. The report lists the results from the current year, as well as those of the two previous years, so we can easily compare the data. The past annual reports are accessible to anyone who wants to consult them on the Language Center website.⁸

PLJ: How did you prepare to do the assessments/evaluation that you use?

Looking back, I see that my interest in this area started very early in my career. I majored in Romance Languages (*Línguas Neolatinas*) in Brazil. From the time I started teaching in private and public secondary schools, I always had an interest in assessment and evaluation of linguistic abilities. I was fascinated by the differences between the speech, reading, and writing of the students from the elite private school where I taught in the morning and that of the working-class village where I taught in the afternoon. What I really wanted was a map, so to speak, that would tell me how people acquire the command of a language, how to measure it, and how to organize language teaching so that both privileged and underprivileged students could develop up to their potential. That is why, after starting an M.A. in Linguistics, I decided to switch to an M.A. in Education. The catalyst for this change was an inspiring workshop that I took with the renowned Brazilian educator and author Magda Soares, who was kind enough to recommend me an excellent bibliography on the matter. The books were predominantly in English, convincing me that I should pursue additional training in the US. I applied for a Ph.D. in the School of Education at Stanford, which was then, like now, the top education school in the country. I studied under Professor Robert Politzer with a concentration in Applied Linguistics, focusing on evaluation. My dissertation explored ways of measuring linguistic and communicative competence among Luso-Americans. This was in the pre-ACTFL era.

⁸ <https://www.stanford.edu/dept/lc/language/about/annualreports.html>

Later, not long after I started teaching Portuguese for Foreign Speakers at UC Berkeley, the OPI and SOPI began to be offered nationally by the Center for Applied Linguistics. I took what I believe was the first-ever SOPI training offered for Portuguese, conducted by Toni Cowles. It was a revelation, because it brought together everything I'd been looking for. That was followed by certification in the OPI a few years later, when I was already at Stanford. My move to Stanford had much to do with knowing that the newly-formed Language Center was grounding its foundations on the ACTFL guidelines. Subsequently, I got my credential in the Writing Proficiency Test. I have been rating both at Stanford and for the Language Testing Institute for 12 years, and I have never stopped learning in this area. Being at the Stanford Language Center has a lot to do with this.

PLJ: Why is that? In what ways has Stanford contributed to your preparation?

Stanford offers a wonderful environment for educators who are interested in evaluation. It actively encourages training and certification through ACTFL in all languages. Sixty-one percent of our full-time instructors (38 people) are certified, and an additional eleven have begun the certification process. Ninety-five percent of all Stanford language instructors (lecturers and graduate students) have participated in the initial stages of oral proficiency training and certification. This is a unique situation, as it is rare in the United States to have even a handful of instructors with such training working together in the same institution. We definitely get an amazing amount of support both from the administration and by exchanging ideas with colleagues.

When I first got my credentials, there had not been an OPI Workshop in Portuguese in the country for almost 10 years, due to low enrollment. Stanford teamed up with ACTFL, covering

some of the costs for the training, and we were able to organize a training workshop for ten instructors of Portuguese from universities such as Brown, Arizona, and Columbia, among others.

The Stanford Language Center not only provides financial support for our training, but also brings ACTFL trainers to organize different kinds of workshops for instructors. This keeps us up-to-date in developments in the field as a group, and fosters further growth through interaction with colleagues at other institutions.

PLJ: In a nutshell, why ACTFL instruments? There are other exams around.

The basic reason for our using the ACTFL framework is very simple: it has it all. It is based on solid standards, presents guidelines that very clearly describe the final result in the various stages of acquisition, uses exhaustively tested instruments, and is supported by an impressive body of research that extends back to its formation as an offshoot of the Interagency Language Roundtable (IRL) scale/interview, which had been making relevant contributions to the field of linguistics since the 1950s. None of the other exams has this track-record or the level of consistency and reliability of these combined tactics. For instance, the Common European Framework dates to 1991 – its body of research cannot yet be compared to the North American.

Isolated national exams of Portuguese, even if communicative (which not all of them are), satisfy their main objectives of awarding a distinction or of designating a certain level of command of Portuguese to allow enrollment in Lusophone national institutions when students are applying for a degree from that institution. But they have yet to offer the solid guidelines or developmental descriptions of proficiency levels, as the IRL/OPI do, and to develop a comparable body of research. In addition, an isolated exam does not allow for inter-language comparisons. As I've

mentioned, at Stanford we can easily tell you which language is working more effectively with higher-level abilities, or compare similarities and differences between the performance of second-year students of several languages, such as Russian, Spanish, and Portuguese, knowing that they have all been evaluated through the very same instrument. This is one, but not the only, argument to justify the fact that the ACTFL/IRL scales and instruments are the measurement standard in the US. The American federal government, private foundations, state board of education, companies, and educational institutions all understand what an Advanced Mid or Level 3 is, be it in Portuguese, Swedish, German, French, or Greek. Any other classification system would not make sense in our context. It is as simple as that. All of these factors make ACTFL a logical choice for us, and for anyone in the United States who is serious about language evaluation.

PLJ: What would you suggest to faculty as a first step to begin using assessment/evaluation in their Portuguese programs?

One can never stress enough that the first and most important step for language programs, courses, and even individual lessons is to identify the target. In other words, one must identify a desired result. This is most effectively done by using a reference framework, which will reflect a program's definition of language and linguistic competence. In other words, objectives will reveal one's beliefs of what language and linguistic competence are. As I mentioned before, at Stanford we use the ACTFL Guidelines as our framework. You don't have to be trained as an ACTFL/OPI-rater to use this framework. Even before I became a trained rater, studying the ACTFL speaking and writing guidelines changed the way I taught and evaluated my students. They really gave me much deeper understanding of how to conduct assessment, and of how I should go about teaching.

When instructors start planning, they often tend to choose individual learning activities that clearly emphasize a certain skill or topic. And they choose activities they know students are going to like – such as a song, a video clip, a game, a field trip. This may engage the students momentarily, but if instructors do not have clearly-defined objectives and sound assessment strategies, as well as clear understanding of the current level of ability of their students, it's likely that such activities will remain disconnected from each other, and may not be very productive in terms of fostering language acquisition.

What I suggest to faculty in terms of evaluation is to use the “backward design”⁹, starting with the most general planning (program or course). The key for the entire process is the definition of the learning goals (What do you want your students to be able to do?)¹⁰, based on the students' current stage of language acquisition (What can they do at this point?). Once those goals are defined, one works “backwards” to determine how to assess whether students have or have not attained the stated goals (How will you know that they have achieved the goal? What instruments will you use to assess them?). The evaluation instruments will have to match the objectives in terms of concrete and observable evidence. Operationalized objectives are also called *rubrics* – the University of Minnesota has an excellent site on this, which I highly recommend.¹¹ Planning activities to help students reach the objectives can only come after these two preliminary stages have been completed. At this stage the instructor should ask himself or

⁹ Genesee, F. and J. Upshur. (1996.) Classroom-based Evaluation in Second Language Education. Cambridge: Cambridge.
Harris, M. and P. McCann. (1994). Assessment. Oxford: Heinemann. Wiggins, G.(2005) Understanding by Design. Alexandria, VA : Association for Supervision and Curriculum Development.

¹⁰ Notice that we are using the communicative approach; therefore, it does not make sense to define the objectives in terms of “What do you want the student to know?” Even if the objective addresses culture, we want the student to be able to do something observable. In other words, the objective should be defined as a rubric, in operational terms. Examples include the use appropriate structures to greet, meet and take leave in their immediate environment; simulate telephone conversations, and activities such as making reservations in restaurant and hotels; express desires, preferences and complaints; present a rehearsed 20 minute report on a cultural festivity in a Lusophone country, etc.

¹¹ http://www.carla.umn.edu/assessment/vac/evaluation/p_4.html.

herself, “What is difficult for the students at this point? What can they not achieve by themselves? What might get in the way of this specific learning objective?”

The backwards planning summarizes quite effectively the essence of teaching and evaluating. If one keeps cycling through the three basic steps of (a) identifying specific learning goals; b) defining acceptable evidence; and c) planning instruction and class activities with reference to students’ progress, it becomes much easier to assess/evaluate the effectiveness of a lesson, a course, or a program. Clearly defining learning objectives is a productive and insightful experience for any administrator or instructor. Moreover, having access to these objectives guides students through self-assessment and learning, as they know what is expected of them. Finally, the results achieved by the students are more revealing of the effectiveness/non-effectiveness of a lesson, course, or program than a questionnaire that solicits students’ opinions. Of course, it’s also important to have students’ evaluations by midterm and at the end of a course, but these data complement one another, and I am a firm believer that concrete results are worth more than opinions. They really represent what the lesson, course, or program has achieved.