

# ***Assessing Foreign Language Proficiency of Undergraduates***

***Richard V. Teschner***  
***Editor***



Heinle & Heinle Publishers  
Boston, Massachusetts 02116, U.S.A.

© Copyright 1991 by Heinle & Heinle. No parts of this publication may be reproduced or transmitted in any form or by any means electronic, or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher.

Manufactured in the United States of America.

Heinle & Heinle Publishers is a division of Wadsworth, Inc.

ISBN 08384-39152

10 9 8 7 6 5 4 3 2 1

6

# Two Philosophies of Language Program and Language Testing Design

*Hector Hammerly*  
*Simon Fraser University*

Two major philosophies of language program design—the holistic approach and what I call the “cumulative mastery” approach—are largely incompatible, as are their testing design correlates. This essay will define and briefly discuss both philosophies and then consider their implications for testing, with examples from oral testing.

## **Holism versus Cumulative Mastery**

The holistic (or global) philosophy holds that a second or foreign language (FL) is best “acquired” as a whole, the way the native language is acquired. Free communication is seen not as an eventual goal but as the essential activity from which grammaticality emerges via negotiation. If there is enough communication, errors will gradually disappear, or so it is claimed. (See, among many others, Krashen & Terrell, 1983; Savignon, 1983; and Swaffar, 1991.) According to this philosophy, FL acquisition in the classroom will be successful if it re-creates the natural communication conditions under which a language is acquired in early childhood or in the field.

For years, communicative syllabi have been opposed to sequential grammatical syllabi (Wilkins, 1976), but I think this is a false dichotomy. A structural sequence can be combined with integration and attention

to communication (Hammerly, 1982). I am not aware of any language program in which what is taught one week is no longer important the next. Moreover, there seem to be few FL programs left in which communication is not (at least in theory) a major goal.

The opposite of holism is the philosophy I call cumulative mastery. This is a concept that can be applied to the guided learning of all complex skills. In the FL field, it means teaching and learning a language step by step, bottom-up, while integrating and using (top-down) what is learned, for gradually freer communication, as it is learned, together with everything taught and learned to that point.

Tests based on the holistic philosophy tend to be holistic too, such as the Oral Proficiency Interview (OPI). On the other hand, tests that agree with the cumulative mastery philosophy concentrate on measuring control and meaningful use of what the students have specifically been taught and have had adequate opportunity to study and master. (This measuring is done first in context-free or limited-context formats and gradually in context-rich and eventually free formats.)

The most highly developed pedagogical manifestation of the holistic philosophy—praised as “most successful” by, for example, Krashen (1984)—is the so-called immersion approach (e.g., Lambert & Tucker, 1972; and Swain & Lapkin, [n.d.]). In this approach, students are not taught the FL as such, but the entire curriculum is (initially) in the FL.<sup>1</sup> The results, though ballyhooed for many years as outstanding, have always been poor, even with the original St. Lambert group whose “success” started the whole movement (Spilka, 1976). The pseudo-immersion approach assumes that one can learn an FL quite well in the classroom while doing something else. Approaches based on the holistic philosophy are misguided in putting the communicative cart before the linguistic horse. While very young children who are surrounded by native speakers and are free of linguistic interference succeed in acquiring their native language “simply” by using it, older children and college students in FL classrooms find themselves under radically different conditions: they already know a language, and they have to share the attention of only one native speaker (if that) with 20 or 25 other students.

Thus, it is naive to even try to reproduce in a FL classroom, crowded with older children or young adults, the process of native language acquisition by young children, which is a once-in-a-lifetime event. Second language acquisition in the field is usually quite successful up to about age 12 or 13, but again, the sociolinguistic environment in a German classroom in Cincinnati is very different from that in the city of Bonn. A FL classroom can be neither a playground nor a foreign country.

Since the immersion approach is the most highly developed manifestation of the holistic philosophy, it should also be its litmus test (Hammerly, 1987). Immersion programs are sociopolitically successful. In Canada, parents have high expectations of French immersion and assume these programs will help their children secure good bilingual positions. Demand has been so great that many parents have stood in line all night to register their children in French immersion. Not only that: parents hear their children speak rapidly, and so they assume they are speaking excellent French. The students themselves have high expectations, and most assume they speak very good French—until further studies or work experience brings them face to face with their ungrammaticality, with the fact that what they speak is really “Frenglish.”

Linguistically, pseudoimmersion has always been a failure. Much of the research supporting this conclusion is summarized in Hammerly (1989), especially chapter 2. Suffice it to refer here to the results of three studies:

- 1) Spilka (1976) found that after six and seven years in French immersion, 40 children in the much-praised original St. Lambert experimental groups had one or more grammatical or vocabulary errors in 52% of their spoken sentences during story retelling. Only about 6% of the sentences of francophone children the same age were incorrect. Further, there was no evidence of progress in accuracy among the immersion pupils from grades 1 to 6, and the gap in correctness between them and francophones increased over the years.
- 2) Pawley (1985) administered Foreign Service Institute interviews to 97 Grade 11 students in early (6,600 hours) and late (3,000 hours) French immersion. The great majority received a rating of 2 or 2+ (on a scale of 0 to 5). In short, it took many years to get a very faulty 2/2+. Compare this to the six months it takes many well-motivated young adults to reach level 3 in speaking in intensive, systematic French or Spanish programs at the Foreign Service Institute and to similar results that can be attained in two or three years in good-quality university language programs.<sup>2</sup>
- 3) Pellerin and Hammerly (1986) reported on interviews with a small number of French immersion students after 13 years in the program. Although most questions in the interviews were easy, on the average these students had one or more grammatical or lexical errors in 53.8% of their simple sentences.<sup>3</sup> Pseudoimmersion evidently doesn't work. It produces not “functional”—as it is claimed—but dysfunctional bilinguals.<sup>4</sup>

The crucial difference between the holistic and the cumulative mastery philosophies lies in how the development of FL competence is viewed. Holists see the whole system as gradually becoming more and more nativelike, as errors slowly disappear.<sup>3</sup> The problem with this view is that one can communicate successfully in a language classroom (if success consists in just conveying meaning) while speaking quite ungrammatically. Important distinctions in languages like Spanish or French are communicatively redundant and thus not the object of negotiation; for example, one can communicate quite successfully in these and other languages without ever mastering (and only rarely having to negotiate) noun gender. When the ability merely to convey messages is stressed, the incentive to communicate grammatically either doesn't develop or is soon lost.

In contrast, cumulative mastery is based on the view that FL competence can be successfully developed in the classroom only when the language system is built up and integrated cumulatively a small part at a time, through careful teaching. Once part A has been taught and learned, it is used to communicate to the extent that it allows; then part B is taught and learned, followed by the meaningful use of A and B together; then part C is taught and learned, followed by controlled communication using A, B, and C, but no more; and so forth. (Of course, as each part is added the relationship between the parts becomes more complex.) Students are discouraged from using partial systems for global ends, and errors are not allowed to become habitual. At the same time, the range and variety of communication gradually increase.

With a cumulative mastery philosophy, students need frequent diagnostic feedback on how well they have learned what they have been taught and are expected to learn. Thus, tests are essential, at least in their diagnostic function. But before we can discuss the design of oral language tests that foster cumulative mastery, we need to consider briefly two points: the proficiency movement and the relationship between language programs and language testing.

## **The Proficiency Movement**

The proficiency movement (PM) has had a mixed effect on FL teaching. Had the PM strongly emphasized control of specific language structures rather than stressing communication so much, its effect on the profession would have been entirely positive. As it is, many programs have abandoned the unseaworthy ship of grammar-translation only to land on the shores of ungrammatical babble.

By calling a rating of 2 on the 5-level ILR scale “advanced” proficiency, ACTFL has institutionalized as normal the poor results that were reported on a quarter-century ago by Carroll (1967). Thus, an opportunity to inspire the profession to improve its performance was lost. It would have been far better to identify those few programs which were producing level 3s and higher, to determine what made them more successful and to make those characteristics known to the rest of the profession.

The PM has also severely distorted the purpose of proficiency interviews, which are at the heart of the ACTFL ratings. The proficiency interview was initially designed—and is still used, at least by the government (as the ILR interview)—to determine whether a person can function in a job that requires proficiency in an FL, without any reference to an instructional program. To determine the limits of the interviewee’s proficiency, interviewers lead interviewees into situations in which they must try to survive linguistically well beyond their control of the language. While this procedure seems desirable to evaluate proficiency for job placement purposes after a language has been learned formally or acquired naturally, a linguistic-survival mentality is unnecessary and indeed linguistically harmful to students within FL programs. It encourages them to try to survive communicatively in interviews—and in general—regardless of grammaticality. The technical difficulties in using such “a very blunt instrument” as the OPI (Meredith, 1990, p. 295) within FL programs might someday be overcome, but the fact that the OPI rewards premature survival skills makes it completely unsuitable for our classrooms.<sup>6</sup>

A high level of proficiency is, of course, the aim of any serious language program; however, the word *proficiency* has become so distorted in its use that I prefer to emphasize the concept and goal of second language competence (*second* being any tongue other than the native language or languages of the student). This is achieved when linguistic competence—structural accuracy, at the core of the model—can be used to communicate (communicative competence) in culturally appropriate ways (cultural competence). (Of course, the PM is a coalition that loosely unites people with very different points of view—those who strongly believe in linguistic accuracy not too comfortably together with those who think that communication is all that matters. And that is the reason the word *proficiency* has lost much of its original meaning, which was strongly oriented toward language structure.) “Teaching for proficiency” à la ACTFL guidelines seems to mean helping students progress from holistic communicative survival in which numerous errors are made, through a long series of holistic stages of communication with fewer and fewer errors, until more or less error-free communication is

achieved. This is a fundamentally incorrect view of how to develop a high level of second language competence in the classroom. ILR interview ratings were never meant to be the basis for an instructional sequence or plan. Some improvement will occur despite a holistic emphasis on communicative survival, but students will soon settle for hard-to-remediate “Spanglish,” “Frenglish,” and so on.

How can communicative fluency and linguistic accuracy be attained? Both can be attained by emphasizing accuracy first, keeping communicative activity under control, and only gradually allowing communication to become freer and more fluent as control of structure expands and by not placing students in the position of making far more errors that can be effectively corrected, and thereby forming poor linguistic habits. The design of tests can support or negate the design of programs, and vice versa. Since students give the impression of directing much effort toward meeting the demands of tests, it is essential that tests match program aims and philosophy at every step.

Tests can contradict the aims of the FL program in a variety of ways. Primary reliance on written tests, for example, negates an oral-emphasis goal, no matter how often this supposed goal is stated. (If the main aim of the program really is to develop oral fluency and accuracy, the listening and speaking skills should always—even at the very advanced level—constitute at least 50% of the course grade.) Fill-the-blank test formats, which allow students to largely ignore even the sentences in which the blanks appear, negate the fact that language occurs in context and should therefore be taught, practiced, used, and tested in the context of, at least, short sentences.

Use of OPIs for a grade—anywhere within or at the end of courses or programs—contradicts the aim of grammaticality, for such interviews virtually guarantee that students will concentrate on surviving communicatively rather than attaining linguistic control. Grammar is no doubt part of the OPI—some might say too much a part of it (Bachman & Savignon, 1986; Kramsch, 1986)—but it is such a general, fuzzy part of the OPI that no specific structures can be mastered by the students in preparation for the interview. Thus, it is bound to be perceived as a holistic communicative evaluation. And, so while OPIs may help make any FL program look impressive, their use within programs makes specific cumulative mastery difficult or impossible to attain.

## Two Oral Progress Tests

I will now discuss two other ways of testing the speaking skill, which I believe remains the weakest aspect of language test design, especially in terms of cumulative mastery.

These two types of oral tests are complementary. Both are guided-output tests; that is, they direct the students to produce oral output that makes use of the sounds, structures, or vocabulary the examiner wants to hear. I am referring to laboratory oral tests and progress (not proficiency) interviews. They both focus on the specifics of FL learning, the more objective laboratory test concentrating more on language rules and elements, the more subjective progress interview allowing some expression of personal meaning. (In "objective" items, only one or a few predetermined answers are acceptable; in "subjective" evaluations, student output can be quite varied.) Colleges and universities that have listen/record laboratories can offer laboratory oral tests; all can offer progress interviews.

### **Laboratory Oral Tests**

In laboratory oral tests (see Appendix A for an example), the students (1) listen to simple directions, models, cues, and so on; (2) are given pauses to think after most cues, and (3) record their output during short pauses. In laboratories with remote or teacher-directed student control, only (3) above is recorded onto the student cassette. This means that from a 20-minute speaking test, an approximately 3-minute tape of student answers would have been recorded, one right after the other, in each booth. Trained teaching assistants, with the help of answer keys, can then score the day's speaking test. In larger programs, several versions of each test may be necessary. Thus, mass testing of the speaking skill is possible.

Laboratory oral tests have four sections. Section 1 involves pronunciation. Here students are asked to record their imitation of a series of short phrases they hear, without being told which particular sounds in the phrases will be evaluated. After the second laboratory oral test in the program, the pronunciation section could be replaced by other activities involving close attention to sounds, such as some forms of dictation. Of course, even after pronunciation ceases to be formally tested, students are held to a high standard of pronunciation on all speaking tests.

Section 2 measures control over a sample of the specific grammatical structures that have been taught, practiced, and used meaningfully since the last test. Structures learned earlier can and should be used too, though not tested directly. The grammar section can take the form of correlation, replacement, transformation, or, when necessary, oral translation frames. While the pronunciation section, if any, is best handled without pauses to think, the grammar section requires pauses both to think before responding and, of course, to record the answers. (My

many years' experience with this type of test shows that with short sentences these pauses need not be longer than about eight seconds.)

Section 3 measures control of a sample of the vocabulary learned since the last oral test. I have found that visuals are unreliable in eliciting specific lexical items and that oral translation of short, simple sentences (following an FL model) works much better. A short model FL sentence with simple structure is presented auditorily with its native-language equivalent, and the FL sentence provides the structural pattern for the testing exercise that follows. In effect, the native-language sentence that cues each frame simply provides the lexical pointers the student needs to modify the FL sentence lexically.

Consonant with the goal of developing not just linguistic but also communicative competence, laboratory oral tests go beyond "mechanical" context-free output to require the students to generate meaningful (though not perhaps personally meaningful) sentences under guidance. This is done in section 4, which consists of directions for FL dialogue between the student and one or more native speakers on tape. Directions are given in the native language so as not to give away words and structures and are presented in an indirect manner so that they merely indicate what idea is to be expressed and cannot be "translated" word for word. The students have to generate their own sentences in the FL, quickly referring mentally, as needed, to previously learned FL sentences, and so forth. Directions have to be such that the student cannot totally ignore what the native speakers on the tape are saying.

Detailed, diagnostically useful feedback should be provided after laboratory oral tests, including points per section.

### **Progress Interviews**

The second type of speaking test based on cumulative mastery is the progress interview (see Appendix B for an example). This may look like any other interview, but the questions are carefully chosen to evaluate control over specific structures, vocabulary, and situational contexts that have been taught and practiced and should have been mastered since the last oral test. A series of alternate questions must be prepared on each point, to help prevent students from learning through the grapevine what questions are being asked.

As a check against inconsistency, it is better (but not essential) that two trained interviewers conduct progress interviews and add up or average the scores. With skillful interviewers, about five or six minutes per student, plus a minute or so for scoring, are enough. Feedback is essential: once students have been evaluated, they should each be given

a chart that points out specific strengths and weaknesses and directs them to sources of remediation.

When facilities and personnel allow it, I have found it best to use both laboratory oral tests and progress interviews within the same school term, since the former deals with specifics and the latter is a little more global. For example, in a regular semester of Spanish as an FL one can give two laboratory tests and one or two interviews, as it is unfair to the students and uncondusive to regular study to have just one make-or-break speaking test at the end of the term. I would suggest that no more than about four or five weeks should pass without one or the other type of speaking test being given; by so doing, the students' attention is drawn consistently and continuously to the importance of oral performance.

The weight given to each section of the laboratory test or to each evaluation criterion for the progress interview is also based on the overall cumulative mastery design. Each aspect of second language competence should be emphasized in testing as it is taught and expected to be learned in the program, with an early focus on linguistic competence gradually shifting to communicative competence, and cultural competence instruction interspersed all along.

Control over sounds is stressed early in a cumulative mastery program. Mastering morphological and syntactic structures is a major concern through the intermediate level (usually the second year in college) and to a lesser extent beyond that level. While vocabulary is learned throughout the program, it would be a mistake to emphasize it early, as such premature lexicalization is linguistically harmful—students whose knowledge of vocabulary exceeds their structural control will still want to use that vocabulary freely, which results in fluent ungrammaticality. Fluency itself should not be given significant weight in oral testing until the advanced or very advanced level, for an early or intermediate focus on fluency is detrimental to linguistic accuracy. (Pseudoimmersion students characteristically speak very rapidly.) “Slow-normal” speed is perfectly acceptable.

A cumulative mastery program must be carefully paced; if it is too fast, students will never quite master anything and will emerge from the program speaking as poor a classroom “pidgin” as those taught with holistic approaches like pseudoimmersion. (And as the language is learned it must be used for communication, or students will emerge, as did many in misguided audiolingual programs, speaking nothing at all.) Most college or university FL programs try to “cover” all of the structure of the language much too quickly—and, as a result, superficially. English-speaking students need at least three, and preferably four, semesters for the once-through-but-thorough cumulative mastery of the

structure of languages like Spanish or French. With faster “coverage” there is no mastery, and in the second and third years students require remediation “reviews” instead of being able to build on what they have learned. (Of course, making the language program semi-intensive or intensive reduces its time frame accordingly, enabling students to take courses taught in the language earlier.)

The construct validity of these two types of speaking tests depends on the validity of the cumulative mastery philosophy. Their content validity derives from making sure that only what has been adequately taught, practiced, and meaningfully used is tested; this content will, of course, have to be established separately for programs and materials that follow different content sequences. The reliability of these two types of tests will have to be established through empirical research.

FL program and test design demand much thought and effort. But it is all worth it when one sees—and hears—young people complete a cumulative mastery language program as transitional bilinguals who can speak (and write) the language at level 3, that is, with accuracy (without poor linguistic habits) and with reasonable fluency. From there, they can move beyond the program—by residing in the foreign setting—to become true bilinguals.

## Notes

1. I say “so-called immersion” because it isn’t immersion at all, as the student is surrounded by 25 other students whose performance in the target language is as poor as his or hers. This is pseudoimmersion. True immersion or, simply, immersion is the situation in which the learner is surrounded by native speakers of the target language. This, depending on the learner’s background, level of language training, and age, either can result in a “terminal 2+” or can be very successful. It seems especially successful after a solid foundation in the FL has been established in the student’s country.
2. My experience includes teaching Spanish for three years at the Foreign Service Institute and serving as a proficiency interviewer for the last half of that time. I have also taught Spanish at the university level for many years. See also charts in Cleveland et al. (1960).
3. A transcript of these interviews appears in Hammerly (1989).
4. Other relevant studies are Adiv (1980), Gustafson (1983), Hamm (1988), Lister (1987), and Totto (1983).
5. The assumption seems to be that one can start with many linguistic “holes” (as in Swiss cheese?), gradually fill the holes, and end up with no holes (as in Gouda?). I claim that with this approach many of the linguistic holes will refuse to disappear and the outcome will almost invariably be unappetizing.
6. Why should FL students be encouraged to survive communicatively at the expense of linguistic competence long before they need to engage in real

communication with native speakers? On the other hand, in my opinion the original ILR interview rating scale should be used for hiring (teachers would have to be at least level 4s to be able to bring their students up to level 3), and it would surely prove invaluable for such research purposes as comparing the outcomes of different programs or methods.

## Works Cited

- Adiv, Ellen. "An Analysis of Second-Language Performance in Two Types of Immersion Programs." *Bulletin of the Canadian Association of Applied Linguistics* 2, ii (1980): 139-52.
- Bachman, Lyle F. & Sandra J. Savignon. "The Evaluation of Communicative Language Proficiency: A Critique of the ACTFL Oral Interview." *Modern Language Journal* 70 (1986): 380-90.
- Carroll, John B. "Foreign Language Proficiency Levels Attained by Language Majors Near Graduation from College." *Foreign Language Annals* 1 (1967): 131-51.
- Cleveland, H., G. J. Mangone & J. C. Adams. *The Overseas Americans*. New York, McGraw-Hill, 1960. (Reproduced in Ringbom, Håkan. *The Role of the First Language in Foreign Language Learning*. Clevedon, Avon, England: Multilingual Matters, 1987: 66-67).
- Gustafson, Rosanna L. "A Comparison of Errors in the Spoken French of Grades Two, Four and Six Pupils Enrolled in a French Immersion Program." Unpublished project, Master of Arts—Teaching of French, Simon Fraser University, 1983.
- Hamm, Christiane. "The ACTFL Oral Proficiency Interview in a Canadian Context: The French Speaking Proficiency of Two Groups of Ontario High-School Graduates." *Foreign Language Annals* 21 (1988): 561-67.
- Hammerly, Hector. *French Immersion: Myths and Reality: A Better Classroom Road to Bilingualism*. Calgary, Alberta: Detselig, 1989.
- \_\_\_\_\_. "The Immersion Approach: Litmus Test of Second Language Acquisition Through Classroom Communication." *Modern Language Journal* 71 (1987): 395-401.
- \_\_\_\_\_. *Synthesis in Second Language Teaching: An Introduction to Linguistics*. Blaine, WA: Second Language Publications, 1982.
- Kramsch, Claire. "From Language Proficiency to Interactional Competence." *Modern Language Journal* 70 (1986): 366-72.
- Krashen, Stephen D. "Immersion: Why It Works and What It Has Taught Us." Ed. H. H. Stern. *The French Immersion Phenomenon*, Special Issue No. 12 (1984) of *Language and Society*: 61-64.
- \_\_\_\_\_. & Tracy C. Terrell. *The Natural Approach: Language Acquisition in the Classroom*. Oxford: Pergamon Press/Alemany Press, 1983.
- Lambert, Wallace E. & G. Richard Tucker. *Bilingual Education of Children: The St. Lambert Experiment*. Rowley, MA: Newbury House, 1972.
- Lister, Roy. "Speaking Immersion." *Canadian Modern Language Review* 43 (1987): 701-17.

- Meredith, R. Alan. "The Oral Proficiency Interview in Real Life: Sharpening the Scale." *Modern Language Journal* 74 (1990): 288-96.
- Pawley, Catherine. "How Bilingual Are French Immersion Students?" *Canadian Modern Language Review* 41 (1985): 865-76.
- Pellerin, Micheline & Hector Hammerly. "L'expression orale après treize ans d'immersion française." *Canadian Modern Language Review* 42 (1986): 592-606.
- Savignon, Sandra J. *Communicative Competence: Theory and Classroom Practice*. Reading, MA: Addison-Wesley, 1983.
- Spilka, Irène V. "Assessment of Second language Performance in Immersion Programs." *Canadian Modern Language Review* 32 (1976): 543-61.
- Swaffar, Janet. "Articulating Learning in High School and College Programs: Holistic Theory in the Foreign Language Curriculum." *Challenges in the 1990s for College Foreign Language Programs: Issues in Language Program Direction*. Ed. Sally Sieloff Magnan. Boston, MA: Heinle & Heinle Publishers, 1991: 27-54.
- Swain, Merrill & Sharon Lapkin. *Evaluating Bilingual Education: A Canadian Case Study*. Clevedon, Avon (England): Multilingual Matters, n.d. [1983?].
- Tatto, Mabel A. "A Comparative Analysis of Grammatical Errors in the Written Code Between Grade Eleven Immersion French and Grade Eleven Core French." Unpublished project, Master of Arts – Teaching of French, Simon Fraser University, 1983.
- Wilkins, D. A. *Notional Syllabuses*. London: Oxford University Press, 1976.

## Appendix A

### Sample Laboratory Oral Test after About 120 Hours of College Spanish

*Note:* Comments to readers of this article appear between brackets. What the student is expected to say appears between slashes. Specific items being scored are underlined. What the student hears appears in quotation marks. The student does not read anything.

[This speaking test would be preceded by a listening comprehension test and followed by a short, simple reading comprehension test, all within 50 minutes.

This type of test, except to some extent for section 4, makes no pretense of being communicative. Its exercises frankly and openly measure oral control over rules and elements of structure in sentence context. Making such a test "more communicative" would destroy its effectiveness in efficiently and objectively evaluating language specifics.]

"SPEAKING TEST AFTER UNIT \_\_\_."

"Section 1: Pronunciation."

[By this point – 120 hours of instruction – there would be no pronunciation section but perhaps some form of dictation. Thus, this is a sample of the format the pronunciation section in early tests would have.]

"After you hear each phrase twice, record once your imitation of it." [Given the possibility of extraneous noise, each item throughout a laboratory oral test should be presented twice. As several responses would greatly complicate scoring, only one – the first, if several are given – is scored; students are informed to that effect.]

[1.] "¿ <u>Q</u> ué bora es?" / ¿ <u>Q</u> ué <u>b</u> ora es?/	0 1 2
[2.] " = " / <u>N</u> o sé, <u>M</u> anuel. /	0 1 2
[3.] " = " / ¿ <u>P</u> or qué no <u>m</u> iras el <u>r</u> eloj?/	0 1 2
[4.] " = " / <u>E</u> stá <u>p</u> arado. /	0 1 2
[5.] " = " / <u>P</u> reguntémoselo a <u>P</u> aco. /	0 1 2
Pronunciation Total	/10

[Ten likely pronunciation trouble spots have been identified in advance and should each be scored on a 0/1 basis – 1 if the sound is pronounced the way a native Spanish speaker would pronounce it, 0 if it isn't. (As there are two trouble spots per phrase, up to 2 points per phrase are possible.) Trying to establish various degrees of pronunciation acceptability for each sound makes the scoring task so slow that it becomes onerous and unmanageable. For the benefit of assistant scorers with limited experience, the answer key should list the most likely substitutions. There need not be a meaningful relationship between the phrases.]

"Section 2: Grammar."

[In the last 20 hours, the following grammatical structures, among others, have been studied: periphrastic future (*ir* + *a* + infinitive), morphology and usage of the two past indicative tenses (preterite and imperfect), and positive and negative constructions using commands and object pronouns.]

"In the following exercises, listen carefully to the instructions and examples; think – don't speak – during the pauses provided for you to think, and record each answer only once when asked to record."

"Exercise 1: Future Construction."

"Change the following sentences from the present tense to the future construction:"

[The periphrastic future might not be communicatively derived from the present indicative, but all that is being determined here is whether

the students can form it. Any tense the students have already mastered would do in the cue sentences, as long as the infinitive is not given to them, for this would make it too easy. The examinees should have to produce the infinitive, as well as the rest of the construction, in their minds.]

“Example: You hear: *Tenemos un examen.* [2] Think; don’t speak yet. [Short pause.] You would change it to (record): [short pause] *Vamos a tener un examen.*”

“Now the exercise begins.”

“1. *Pepe conversa con ella.*” [2]

[Pause to think.]

“Now record.” [Pause.] [Same instructions throughout.]

*/Pepe va a conversar con ella./* 0 1

(Etc.)

[Scoring: the underlined part of each sentence must be correct for the student to get any credit. A full point is given when the whole sentence is correct; a half-point, if there are one or more grammatical or lexical errors or omissions outside the underlined part.]

“Exercise 2: Preterite.”

“Change the following sentences to the preterite.”

[Examples, cues, and pauses as above to produce five test sentences such as these:]

“1. *No venden los muebles.*”

*/No vendieron los muebles./* 0 1

(Etc.)

“Exercise 3: Commands.”

“Change the following sentences into commands, substituting pronouns for nouns.”

[Examples, cues, and pauses as above to produce five sentences such as:]

“1. *Ud. quiere dejar el auto.*”

*/Déjelo (Ud.)/* 0 1

(Etc.)

“Section 3: Vocabulary.”

[All of the grammatical structures used in the vocabulary exercises are thoroughly known to the students. Only their ability to produce some of the most recently learned lexical items is being tested.]

“Translate the following sentences, using as a model the structure of the latest Spanish sentence you hear.”

“Example: Repeat: *Vamos a leer un libro.* This means: We’re going to read a book. [2] Now translate: We’re going to buy a record player.”

[Pause to think.] "You would say: [pause to respond] *Vamos a comprar un tocadiscos.*" [Note that the correct answer in Spanish is given as the model for the next sentence.]

"Exercise 1."

"Repeat: *Acaba de salir.*" /*Acaba de salir.*/

"She's just left. *Acaba de salir.*" [2]

[1.] "Now translate: He's just lied again." [2]

[Pause to think.] Record."

/*Acaba de mentir otra vez.*/

0 1

"Repeat: *Acaba de mentir otra vez.*"

(Etc.)

"Exercise 2."

[Directions, etc., as in exercise 1 above. Model sentence: *¡Qué viaje tan lindo!* = What a beautiful trip!]

"*¡Qué viaje tan lindo!*"

[1.] "What an expensive gift!"

/*¡Qué regalo tan caro!*/

0 1 2

(Etc.)

"Section 4: Dialogue"

[This dialogue calls on the students to use meaningfully, in an integrated manner, what they have learned so far, with emphasis on the past tenses.]

"In this part you will participate in a guided conversation with *los señores Iglesias*, whose voices are on tape. Follow the usual directions for pauses to think and to record."

(*Sr. Iglesias*): "...y, *¿qué le parece San Antonio?*"

"Answer that you like it very much; add that you especially like the parks, the restaurants, and the Spanish colonial buildings." [2] [Fairly long pause to think.]

"Record." [Pause to record; about 12-15 seconds are needed to record a rather long sentence like this one.]

/*Me gusta mucho; especialmente me gustan los parques, los restaurantes, y los edificios coloniales.*/

0 1 2 3

(*Sra. Iglesias*): "*Dígame . . . ¿es usted del este de los Estados Unidos?*"

"Answer that you're from Wisconsin. Add that when you were a child you lived in the country."

/*Soy de Wisconsin. Cuando era niño(a) vivía en el campo.*/

0 1 2 3

"But add that you were a little bored and tired of the cold."

/*Pero estaba un poco aburrido(a) y cansado(a) del frío.*/

0 1 2 3

“So explain that when there was an opportunity for a good job, you came here.”

*/(Así que, De modo que) cuando hubo (una) oportunidad de un buen empleo (trabajo), (me) vine aquí./* 0 1 2 3

*(Sr. Iglesias:) “Y, ¿de dónde cree que es mi esposa?”*

“Say you think she’s Chilean, but you’re not sure.”

*/Creo que es chilena, pero no estoy seguro(a)./* 0 1 2 3  
/15

[Scoring is 3 points for a complete sentence without errors or omissions; 2 points if it has 1 error; 1 point if it has 2 or 3 errors; and 0 points for more than 3 errors or no response. Alternate forms considered correct in educated standard Spanish are accepted.]

#### Test Totals

Listening Comprehension: \_\_\_ /30

Speaking ability:

Pronunciation: \_\_\_ /10

Grammar: \_\_\_ /15

Vocabulary: \_\_\_ /10

Dialogue: \_\_\_ /15

Subtotal: \_\_\_ = \_\_\_ /50

Reading comprehension: /20

Overall test total: \_\_\_ /100

Other, shorter tests involving listening and reading comprehension and simple written tasks can be administered between the longer tests and/or interviews that emphasize speaking ability.

## Appendix B

### Sample Progress Interview after About 120 Hours of College Spanish

By carefully selecting questions in advance and improvising a little, interviewers can get students to use those grammatical structures whose control needs to be evaluated (because they have recently been taught, practiced, etc.) at each point in the program. This plan doesn’t always work, however, for unlike the directed dialogue just discussed, the more open-ended questions of an interview give students a chance to display their keen ingenuity for talking around the structures they should have mastered but have not. After seeking to put the student at ease in English, if necessary, the interviewers would ask a total of five or six questions of the sort exemplified below, plus a few short follow-up

questions as needed. (Four or five alternative questions for each of the questions listed would be used with different students.)

(1. *Gustar, parecer, and present tense in general:*)

“¿Le gusta más el campo o la ciudad?”

(Follow up:) “¿Por qué?”

(2. *Commands:*)

“Deme instrucciones en cuanto a qué hacer:”

“Tengo hambre. ¿Qué bago?”

[E.g., /Vaya a la cafetería, etc./.]

(Follow up:) “Y entonces, ¿qué bago?”

(3. *Periphrastic Future:*)

“¿Qué va a hacer Ud. el fin de semana?”

(Follow up/expansion:) “¿Y sus amigos/amigas?”

(Follow up:) “¿Y entonces?”

(4. *Preterite and Imperfect:*)

“Díganos qué hizo, etc., esta mañana desde que salió de su casa hasta que llegó aquí.”

(Be sure to follow up with questions that require the student to shift tense aspect, such as “¿Qué pasó entonces?” “¿Cómo era(n)?” “Y entonces, ¿qué hizo?” “¿Puede describirlo/la?”)

[5. A topical question for a longer answer could also be asked.]

While the above interview is in progress, the two interviewers should each be marking a criteria sheet cum structure list such as the one illustrated below. The students know quite well that they are being tested; still, it helps them if these sheets are not filled out in full view (the sheets can be kept, for instance, inside the covers of large books). At the end of the interview, each interviewer should have decided on one point for each criterion line and an indication (e.g., numbers or check marks, agreed on in advance) of which of the structures on the list the student needs to work on further.

**Sample Progress Interview Evaluation Sheet**

Course: SPAN 101 (second semester)

Name:

Date:

Comprehension 0 2 4 / 6 8 10: \_\_\_\_\_

Pronunciation 0 2 4 / 6 8 10: \_\_\_\_\_

Grammar 0 5 10 / 15 20 25: \_\_\_\_\_

Vocabulary 0 2 4 / 6 8 10: \_\_\_\_\_

Fluency 0 1 2 / 3 4 5: \_\_\_\_\_

Subtotal 1: \_\_\_\_\_/60

Subtotal 2: \_\_\_\_\_/60

Total: \_\_\_\_\_/120



## Structure Checklist

## Comments and/or Examples

---

Present, irregular forms  
 - - - - - , usage  
 Gender/number agreement  
 .....(Etc.).....  
 Imperfect, forms  
 Preterite, regular forms  
 - - - - - , irregular forms  
 Imperfect and preterite, usage

(Only the structures emphasized since the last oral test or progress interview and those which tend to present ongoing difficulties need be listed.)

The criteria given above would be weighted differently at different points in the program, in accordance with the program's teaching philosophy. Students should be informed in advance of what the various criteria are and how they will be weighted.

Each question should not be asked more than twice at a slow speed within normal range. Of course, if a student doesn't understand a question, he or she fails that portion of the interview. Other than that, comprehension need not be emphasized within progress interviews provided that it is measured often and in other ways, that is, via listening and reading comprehension activities and tests.