

Test de opción múltiple, redacciones dirigidas y de tema libre: ¿sirven para medir el progreso de nuestros alumnos?¹

(Multiple-Choice Tests, Controlled and Free Writing: Can They Be Used to Measure Our Students' Progress?)

Asunción Martínez-Arbelaiz²

University Studies Abroad Consortium

Resumen

Los profesionales de la enseñanza de lenguas extranjeras o segundas lenguas se preguntan si sus alumnos progresan en el continuum del dominio de la lengua en cuestión. Sin embargo, dar evidencia «objetiva» del progreso de los aprendices nunca es tarea fácil. En USAC, un consorcio de universidades de Estados Unidos que ofrece cursos de español en países de habla hispana, se han adoptado diferentes formas de evaluación (*assessment*), como el test de opción múltiple, las redacciones dirigidas y las redacciones libres para diagnosticar el desarrollo lingüístico de los alumnos; todo ello tiene un impacto directo en el diseño de nuestros cursos.

Abstract

Professionals in the teaching of foreign or second languages often wonder whether their students are making progress in the language proficiency

¹ Elaborado con base en la ponencia presentada por la autora en el IV Congreso Internacional de Lingüística Aplicada, llevado a cabo en mayo de 2013, en el Campus Omar Dengo, de la Universidad Nacional de Costa Rica. Recibido: 19 de mayo de 2012; aceptado: 30 de agosto de 2012.

² Correo electrónico: asuncion@usac.unr.edu

continuum. However, providing “objective” evidence of learners’ development is never an easy task. In USAC, a consortium of universities in the United States which offers courses of Spanish in Spanish-speaking countries, we have adopted several assessment measures, such as multiple-choice tests, and free and directed writing. This evidence is very valuable in order to diagnose learners’ linguistic development; and it has a direct impact on the design of our courses.

Palabras clave: evaluación, desarrollo lingüístico, test de opción múltiple, escritura libre

Keywords: assessment/evaluation, linguistic development, multiple-choice test, free writing

Introducción

La experiencia de estudiar en un país extranjero significa mucho más que puras ganancias lingüísticas, incidiendo en la propia identidad de las personas³, pero los que se dedican a la enseñanza de segundas lenguas tenemos se preguntan si los alumnos progresan y avanzan en el continuum del dominio de la lengua en cuestión. Sin embargo, proporcionar datos del progreso de los alumnos es una tarea más difícil de lo que podría parecer a primera vista. En USAC (University Studies Abroad Consortium)⁴, se llevan años reflexionando sobre este tema y diseñando instrumentos van para medir el desarrollo lingüístico de ese constructo cambiante llamado «interlengua», definido como un sistema lingüístico complejo en desarrollo y con errores también sistemáticos⁵.

En los últimos años hemos experimentado con diferentes formas de evaluación (inglés, *assessment*) en el programa de español como lengua extranjera, de forma que en las líneas que siguen detallaremos las inherentes ventajas y desventajas de los métodos adoptados. Ni que decir tiene que el proceso de evaluación de los resultados de un

³ Phil Benson, Gary Barkhuizen, Peter Bodycott y Jill Brown, «Study abroad and the development of second language identities», *Applied Linguistics Review* III-1 (2012): 173-193.

⁴ Para más información sobre los cursos de lengua que ofrecemos, se puede visitar la página web del Consorcio de universidades: <www.usac.unr.edu>.

⁵ S. Pit Corder, «The Significance of Learners’ Errors», *International Review of Applied Linguistics* V (1967): 161-170; Larry Selinker, «Interlanguage», *International Review of Applied Linguistics* IX (1972): 67-92.

programa completo de español es largo y complejo y nunca va a ser perfecto, pero, siguiendo la línea de trabajo iniciada en Norris, Davis, Sinicrope y Watanabe⁶, en el presente estudio detallamos y resumimos los éxitos y fracasos de este complicado proceso con objeto de contribuir a una mayor transparencia sobre las posibles formas de la evaluación de un programa de lengua en situación de inmersión.

Desde 2003 venimos realizando pruebas sistemáticas en los nuestros cursos de lengua, con fin de saber más sobre aquellas áreas que experimentan un desarrollo y aquellas que no. Durante una serie de semestres consecutivos los alumnos de diferentes programas de español completaron un test de opción múltiple al principio y al final del curso. Comparaciones estadísticas mostraron un patrón de mejora en todos los niveles. Tras un semestre en el que se les pidió un test más reducido y una redacción, lo cual nos pareció excesivamente largo y tedioso para los alumnos, nos limitamos a pedir una redacción al principio y al final del curso. Esta tarea escrita nos dio pistas de áreas o puntos gramaticales en los que los alumnos habían mejorado y aquellos aspectos que todavía requerían más instrucción y atención. Concluimos que todas estas fuentes de información son muy valiosas para diagnosticar el tipo de desarrollo lingüístico de nuestros alumnos y deben tener un impacto directo en el diseño de nuestros cursos. Sin embargo, diferentes formas de evaluar los resultados de aprendizaje nos iluminan diferentes aspectos de la interlengua de los alumnos, por lo que proponemos alternar y variar los métodos de evaluación para poder así llegar a una panorámica más global del desarrollo lingüístico de los alumnos en el extranjero.

¿En qué consiste la evaluación de un programa de lengua?

En inglés existen dos palabras, relacionadas pero diferenciadas: *assessment* y *evaluation*. Según Bachman y Palmer⁷,

⁶ John M. Norris, John McE. Davis, Castle Sinicrope y Yukiko Watanabe, *Toward Useful Program Evaluation in College Foreign Language Education* (Honolulu, HI: National Foreign Language Resource Center-University of Hawai'i, 2009).

⁷ Lyle F. Bachman y Adrian S. Palmer, *Language Assessment in Practice* (Oxford: Oxford University Press, 2010).

[a]ssessment is the process of collecting information about something that we're interested in, according to procedures that are systematic and substantively grounded (...). Evaluation is one possible use of assessment; it involves making value judgments and decisions on the basis of information, and gathering information to inform such decisions is the *primary* purpose for which language assessments are used. We frequently use assessments for evaluation in educational programs, where we may use assessments to identify students' areas of strength and weakness to help them make decisions to improve or facilitate their learning, to select and place individuals into instructional programs, or to decide which students pass a course.

En cuanto a la relación entre *assessment* y *evaluation*, la primera sería la recogida de datos sistemática, y el uso e interpretación que hacemos de dichos datos sería la evaluación, la cual tiene a su vez consecuencias sobre instituciones, programas e individuos. En español, sin embargo, no contamos con dos palabras diferenciadas por lo que utilizaremos la expresión «recogida de datos sistemática» para *assessment* y el cognado «evaluación» para *evaluation*.

Contextualización: ¿por qué debemos evaluar un programa?

Los profesores estamos habituados a evaluar a los alumnos porque se pide dar una nota de su rendimiento al final de cada curso impartido. Para ello dicha nota los profesores tenemos un criterio implícito de lo que se espera que el alumno haga en un determinado nivel y diseñamos pruebas u exámenes orales y escritos que miden si se cumplen esos objetivos o no. Tal nota es fundamental para que un alumno siga subiendo en los diferentes niveles o escalas de competencia de la lengua. En USAC distinguimos cuatro niveles, los cuales se corresponden, grosso modo, con los años que un alumno suele estudiar en una universidad estadounidense. Así, el nivel 1 será el primer año, el cual a su vez consta de dos semestres, el nivel 2 será el segundo año, el nivel tres el tercer año de estudio de la lengua

meta y el cuarto nivel será a partir del cuarto año. Siempre partiendo de un posible alumno que no haya estudiado español nunca o que parta de cero.

En caso de que el alumno ya cuente con algún conocimiento de español al llegar a un curso de español, dicho alumno suele realizar un examen de clasificación por niveles. Según la nota obtenida en un test de opción múltiple, junto con una entrevista oral con uno de los profesores se determinará en qué nivel tiene que estudiar el español el alumno en cuestión.

Sin embargo, en USAC no hay un examen o test de salida estándar, el cual nos determine el nivel con el que salen los alumnos, más allá del examen que cada profesor haya diseñado. Así pues, nuestro primer plan fue que los alumnos volvieran a realizar el test de clasificación que realizaron al llegar al programa pero esta vez al final de un semestre de estudio intensivo de español. El interés de proporcionar evidencia de lo que aprenden los alumnos tiene un doble propósito: por un lado realizar una autocrítica del programa con su correspondiente actuación de mejora y, por otro lado, ofrecer datos concretos sobre lo que estamos haciendo a las universidades que nos envían estudiantes todos los semestres (*accountability*).

Aplicación de test de opción múltiple

En 2003 se empezaron a medir los resultados de los cursos. Si se quiere saber cuánto aprenden nuestros alumnos, una manera de «medir» el progreso podría ser volver a pasar el mismo test de clasificación pero una vez terminado el semestre. Todos los programas utilizan el mismo test de 85 preguntas de elección múltiple para clasificar a los alumnos por niveles. Este test de clasificación por niveles con el que ya contábamos no sigue un orden de preguntas más fáciles o difíciles sino que los ítems están mezclados en cuanto a su grado de dificultad. De esta forma, se pide a los alumnos que completen todas las preguntas que puedan. El test en cuestión incluye preguntas de morfología verbal (Ex. 1 y 2), de expresiones

particulares (Ex. 3 y 4) o de léxico, incluyendo frases hechas, como se puede apreciar en los ítems 5 y 6:

1. Juan y yo...en la universidad.
a) estudiáis b) estudiamos
c) estudia d) estudio
2. María está... el piso.
a) arreglado b) a arreglar
c) en arreglar d) arreglando
- 3., no merecía tal castigo.
a) Era quien era b) Fuera quien fuera
c) Será quien será d) Sería quien sería
4. «... bicicletas a tres euros la hora», decía el letrero.
a) Ud. alquila b) Se alquilan
c) Uno alquila d) Él alquila
5. La ventana está.....
a) ancha b) cerrada
c) agradable d) grande
6. «¿ A qué se dedica usted?». Significa:.....
a) ¿En qué piensa usted? b) ¿Para qué se prepara usted?
c) ¿Qué hace usted? d) ¿A qué aspira usted?

Salvo los alumnos de nivel 1, los de los demás niveles contaban con una nota para comparar su nota de entrada y de salida. La correspondencia entre la puntuación y los niveles de USAC es la siguiente: de 0 a 25 puntos, nivel 1; de 25 a 42, nivel 2; de 43 a 64, nivel 3. Finalmente aquellos alumnos que tenían una puntuación superior a 65 puntos podían tomar el nivel 4.

El test de nivel se combina con una entrevista oral, pero es bastante raro que un alumno obtenga más de 64 puntos, a no ser que sea

un hablante casi nativo, por lo que se aceptan alumnos con una puntuación algo más baja en el nivel 4. Como se puede ver en la Tabla 1, en todos los niveles la nota media de salida fue más alta, aunque conforme sube el nivel es más difícil ver el progreso.

Tabla 1. Resumen de los resultados de las pruebas de nivel por semestre

Nivel	Inicio (Sept. 2003)	Fin (Dic. 2003)	Inicio (Enero 2004)	Fin (Mayo 2004)	Inicio (Sept. 2004)	Fin (Dic. 2004)	Inicio (Enero 2005)	Fin (Mayo 2005)
1		46,75(30)		42,38(44)		43,16(31)		45,11(34)
2	33,33(41) ¹	51,38(42)	32,63(40)	51,75(46)	33,04(66)	53,58(61)	33,83(46)	52,14(49)
3	45,93(38)	57,86(41)	44(47)	57,2(65)	47,62(40)	58,74(44)	42,51(40)	55,74(48)
4	66,22(13)	72,32(16)	52,18(20)	61,6(26)	64,2(22)	67,95(25)	55,87(11)	62(34)

El final del nivel 1 está todos los semestres por encima de 42 puntos, es decir en el principio del curso de Composición 1, por lo que podemos concluir que el principio de la adquisición se produce de forma rápida. El nivel 2 comienza con un Intermedio I alto y termina en el final de Composición 1. El nivel 3 comienza en Composición I más bien bajo y termina en un curso más, Composición II (54-64). Finalmente, se observa una ralentización del progreso en los niveles superiores que está en consonancia con las representaciones de las escalas en forma de cono invertido utilizadas, por ejemplo, en el Marco Común Europeo de Referencia. En la representación del dominio total de un idioma, las primeras escalas o niveles contiene menos contenidos y menos exigencia que las últimas. Por lo tanto, se puede concluir que en los niveles más bajos se requiere menos tiempo para progresar, mientras que en los niveles avanzados el progreso es más difícil de conseguir, y por lo tanto de mostrar.

La prueba de nivel de tipo test fue fácil de administrar, de corregir y de interpretar. Durante cuatro semestres consecutivos la información obtenida fue muy similar y tendía a repetirse, por lo que decidimos no continuar acumulando datos de la misma naturaleza. Un alumno que toma nuestros cursos sin haber realizado cursos universitarios de

¹ El número en paréntesis indica el número de alumnos que hicieron la prueba.

español previos puede llegar, tras un semestre de estudio, a nivel de composición, es decir, con un español de tercer año de universidad. Asimismo, un alumno de nivel dos puede acceder a los cursos superiores, si toma las clases con seriedad. Sin embargo, vimos que el progreso es más limitado, por lo menos en lo referente al conocimiento de la gramática y el léxico que se incluyó en el test, para los niveles más altos.

Redacciones de tema dirigido

Después de esta experiencia que mostró ciertos rasgos de la eficacia de nuestros programas, pensamos que queríamos seguir indagando y buscando información más precisa que la que un test nos puede aportar. Seguíamos con dudas sobre qué aspectos gramaticales y léxicos se desarrollaban más o menos. Así que decidimos cambiar el instrumento de obtención de datos, optando por un sistema que nos diera más pistas sobre los aspectos de la competencia lingüística de los alumnos que mejoraran o que no en los diferentes niveles. Para ello decidimos que los alumnos escribieran una redacción sobre una vacación al comienzo del semestre y que la repitieran al final del curso en cada una de las clases de los programas de España. El estímulo que se les dio para la producción escrita fue el siguiente:

Seguro que recuerdas algunas vacaciones especiales o algún viaje que hayas hecho. Explicanos en unas 250 palabras cómo fueron esas vacaciones o ese viaje. Explica con quién fuiste, por qué elegiste ese lugar, en qué medio de transporte viajaste, dónde te alojaste, qué actividades realizabas, etc. Intenta recordar todos los detalles que puedas. También puedes contar alguna anécdota o si te ocurrió algún pequeño desastre o algo especialmente bueno.

Un total de noventa y ocho alumnos completaron las dos composiciones requeridas. Se les pidió a los profesores que no ayudaran a los alumnos con la prueba de escritura y que no les dejaran utilizar diccionarios, gramáticas u otros materiales electrónicos. Como suele ser

usual en este tipo de investigación, algunos alumnos sólo completaron la redacción del principio o la del final, por lo que esas redacciones se eliminaron del estudio. El cómputo total de alumnos que completaron las dos redacciones en cada nivel fue el siguiente: en nivel 2 fueron treinta y ocho los alumnos que completaron las dos redacciones, en el nivel 3 fueron también treinta y ocho alumnos y en el nivel 4 hubo veintidós alumnos que realizaron las dos tareas de escritura.

Los datos de cada una de las variables mencionadas se insertaron en un programa informático SPSS para elaborar estadísticas por medio de pruebas t para la igualdad de medias. La diferencia de las siguientes variables dependientes resultó estadísticamente significativa entre el principio y el final del semestre en los tres niveles escudriñados:

En el nivel 2:

- Número de palabras
- Cociente entre el número de palabras y las unidades t
- Número de pretéritos imperfectos
- Número de pluscuamperfectos
- Usos del verbo *ser*
- Influencia de la L1

En el nivel 3:

- Número de palabras
- Número de subordinadas incorrectas
- Tipos de subordinadas
- Número de pretéritos indefinidos incorrectos
- Número de pretéritos imperfectos
- Número de pluscuamperfectos
- Usos de tiempos incorrectos
- *Ser* incorrectos
- Concordancia

En el nivel 4:

- Ninguna

No contamos con el nivel inicial porque obviamente no podían escribir nada al empezar. Sin embargo, recogimos las redacciones de este grupo del final del semestre de instrucción pero solo las podemos comparar con las del inicio del nivel 3, algo que dejamos para futuros proyectos. Lo que llama poderosamente la atención es que todas las variables analizadas quedaron estables al compararlas en el inicio y en el final del nivel 4, a diferencia de lo que encontramos en los niveles 2 y 3. En 2004, ofrecimos una descripción más detallada de la codificación, análisis e interpretación de estos resultados⁹. En conclusión, una vez más, el nivel 4 se caracteriza por la falta de progreso observable en la interlengua de los alumnos. A la luz de estos resultados decidimos concentrarnos en este nivel avanzado con instrumentos de recogida de datos diferentes.

Redacciones de otros temas diferentes

Una presunta inferencia de estos resultados, proporcionada por un profesor implicado en una de nuestras reuniones fue la siguiente: podría ser que la tarea solicitada fuera demasiado fácil para los alumnos de nivel 4, es decir, de nivel más alto, por lo que se obtuvo un efecto «techo» (*ceiling effect*). Si la tarea solicitada es demasiado sencilla, ésta no documenta progreso y los resultados se mantienen iguales. Se pensó en administrar un instrumento con otro tema para los alumnos de nivel más alto. Así, en la primavera 2006 diecinueve alumnos de los cuatro programas completaron la misma redacción antes y después del semestre. Se diseñó otra tarea más compleja que pedía que los alumnos argumentaran a favor o en contra de la idea de que el internet favorece el aislamiento. Cambiamos el tema y el tipo de texto requerido. En lugar de una narración, los alumnos tenían que redactar un texto argumentativo, lo cual a priori, parece una tarea más difícil de completar y que requiere más dominio de la lengua meta.

En efecto, comparando los resultados de las redacciones de principio y final de semestre, se documentaron los siguientes

⁹ Asunción Martínez Arbelaiz, «Índices de progreso en la producción escrita de estudiantes de español en situación de inmersión», *RAEL: Revista Electrónica de Lingüística Aplicada* III (2004): 115-145.

cambios: la fluidez aumenta, en lo que respecta a número de palabras, número de unidades t¹⁰, número de palabras en cada unidad t, número de oraciones. Sin embargo, no encontramos una diferencia significativa en el ratio de cláusulas en cada unidad t. Curiosamente, hubo más nexos y de más tipos diferentes.

En lo que respecta a la corrección gramatical, vimos que el uso de subjuntivo sube (la media en la primera redacción fue de 1,31 subjuntivos y en la última de 2,63); pero esta diferencia no fue estadísticamente significativa. Tampoco lo fue el número de subjuntivos incorrectos ni el número de contextos que requieren el uso de subjuntivo¹¹. Ofrecemos las medias de estas variables en la Tabla 2:

Tabla 2. Resultados en el nivel 4

	Media pre	Media post	Valor-p
Palabras	231,73	327,36	0,0004
Unidades t	19,78	23,89	0,04
Palabras/Unidad t	11,83	13,93	0,003
Cláusulas	33,42	43,47	0.008
Cláusulas/Unidad t	1,66	1,79	---
Nexos	11,94	17,89	0,001
Tipos de nexos	3,6	5,6	0,0001
Subjuntivos	1,31	2,63	0,12
Subj. incorrectos	1,0	0,94	--
Contextos	1,63	2,94	--

¿Qué interpretación podemos dar a estos resultados en combinación con los obtenidos en las secciones anteriores? Definitivamente es en el nivel 4, el más alto en nuestra escala de niveles de dominio, donde resulta más difícil documentar el progreso de los alumnos. Esto estaría

¹⁰ Kellogg W. Hunt, «Recent Measures in Syntactic Development», *Elementary English* XLIII (1966): 732-39.

¹¹ Isabelli Casilde A. y Chiyo Nishida, «Development of the Spanish Subjunctive in a Nine-Month Study-Abroad Setting», en David Eddington, ed., *Selected Proceedings of the 6th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages* (Somerville, MA: Cascadilla Proceedings Project, 2005), 78-91. Disponible en: <<http://www.lingref.com/cpp/casp/6/index.html>>.

en consonancia con la afirmación de que es necesario cada vez más tiempo dedicado al aprendizaje para ir subiendo en esta escala potencial señalada en los niveles del Marco Común Europeo de Referencia. Otra interpretación que barajamos fue que nuestro diseño curricular de los cursos avanzados pudiera ser modificado para responder a las necesidades lingüísticas de alumnos de este nivel. Así, a raíz de estos resultados, hicimos una remodelación de los niveles avanzados y decidimos incluir un proyecto, además de una lectura, revisando asimismo los libros de texto adoptados. Por razones de espacio no detallamos aquí evaluaciones posteriores que hemos realizado en algunas de las clases de nivel avanzado, pero podemos simplemente adelantar que nos hemos centrado en documentar el desarrollo de la riqueza léxica.

Conclusiones

Las pruebas tipo test resultaron fáciles de administrar y de corregir pero, como hemos comprobado, proporcionaron escasa información sobre las áreas específicas en las que mejora o se estanca el alumno de segundas lenguas. Los resultados de estas pruebas documentaron la mejora en todos los niveles, ofreciendo un número más alto en cada nivel, pero no arrojaron luz sobre si los alumnos fallan en morfología verbal, en léxico, en complejidad sintáctica u otros aspectos lingüísticos. Sin duda resultaron útiles para documentar la mejora global de los alumnos, pudiéndose establecer comparaciones estadísticas claras tanto comparando la nota del principio y del final de cada curso, como estableciendo comparaciones entre el principio del nivel tres y el final de nivel 1, por ejemplo. Este tipo de comparativa nos proporciona información sobre el nivel final obtenido en nuestros cursos en comparación con el nivel de los alumnos preparados en otras instituciones.

Por otro lado, la escritura, bien sea libre bien sea sobre un mismo tema, nos dio información valiosa sobre lo que el alumno es capaz de hacer al entrar en el programa y lo que puede hacer después de haber estado viviendo rodeado de la lengua meta y asistiendo a clases. En contrapartida, las redacciones resultaron más opacas a la

hora de ofrecer resultados, debiendo pasarse por un tedioso y largo proceso de codificación, siempre sujeto a debate y deliberación.

A pesar de estas dificultades iniciales, una conclusión de estos estudios es que dependiendo del tema asignado, es decir, del estímulo de la redacción, los resultados de la evaluación pueden variar sustancialmente. En otras palabras: la tarea determina o influye en los resultados. Teniendo en cuenta el efecto de la tarea, decidimos en los últimos semestres solicitar a los profesores que les pidan a los alumnos una tarea escrita de entrada y otra de salida en cada nivel, sin especificar el tema. Dejamos por lo tanto en manos de cada profesor diseñar un estímulo de escritura acorde con los alumnos de cada curso. Es esta la línea que pensamos seguir adoptando en el futuro: convencer al profesorado de que una redacción escrita al término del programa da evidencia de lo que se ha conseguido con cada uno de los alumnos, y a la vez, es posible hacer una evaluación del grupo en general, identificando carencias que deberán ser abordadas en cursos superiores.

Las redacciones abiertas en cuanto al tema no permiten establecer comparaciones estadísticas, como las que se han detallado anteriormente, pero nos dan una idea global rápida de lo conseguido por el alumno. Podemos realizar un análisis más cualitativo que nos dé evidencia no sólo de los aspectos lingüísticos que mejoran, sino también de las experiencias vividas por los alumnos y de su grado de integración. Para ilustrar este análisis cualitativo, podemos tomar como ejemplo una redacción de un alumno estadounidense después de pasar un semestre en Puntarenas, Costa Rica. Al leer el texto olvidamos los errores y, evidentemente, el autor satisfactoriamente comunica y contagia la emoción que la experiencia le suscita. Un análisis cualitativo de los temas abordados por los alumnos también puede arrojar luz sobre la eficacia de los programas en cambiar la perspectiva de los alumnos hacia la lengua y la cultura meta, haciéndola, como se puede leer en la siguiente redacción, más matizada y sofisticada:

Antes de mi viaje a Costa Rica me había preocupado acerca de muchas cosas. Me había preocupado de que mi familia no fuera agradable y que no gustara la comida. Pero me equivoqué. Mi familia es

muy entusiasmada por muchas cosas. Yo esperaba que mis profesores me enseñaran mucho español y del océano. Conseguí mi deseo. Estoy aprendiendo mucho español cada día. He aprendido mucho en Costa Rica, he aprendido sobre América Latina, de la naturaleza, y del mundo. Aprendí que hay un mundo fuera de los Estados Unidos y que América Latina tiene una cultura que merece respeto. Pensaba que las culturas de América Latina eran iguales pero aprendí que los países de América Latina tienen una cultura muy diferente.

Aunque no existe una fórmula de evaluación perfecta, el mero hecho de cuestionar, de pensar y de diseñar herramientas que nos permitan valorar lo que conseguimos con cada grupo de alumnos es un paso que todo programa que quiera mejorar debería dar. En estas líneas hemos desbrozado los diferentes caminos tomados en los años que llevamos evaluando nuestros propios resultados y esperamos que sirvan de guía y de inspiración a otros programas de lengua que estén pensando en evaluar sus resultados con objeto de mejorar la instrucción ofrecida.