



Íkala, revista de lenguaje y cultura

ISSN: 0123-3432

revistaikala@udea.edu.co

Universidad de Antioquia

Colombia

Arias Toro, Clara Inés; Maturana Patarroyo, Liliana María
Evaluación en lenguas extranjeras: discursos y prácticas
Íkala, revista de lenguaje y cultura, vol. 10, núm. 16, enero-diciembre, 2005, pp. 63-91
Universidad de Antioquia
Medellín, Colombia

Disponible en: <http://www.redalyc.org/articulo.oa?id=255020409003>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica

Red de Revistas Científicas de América Latina, el Caribe, España y Portugal

Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

Evaluación en lenguas extranjeras: discursos y prácticas* 1

Mg. Clara Inés Arias Toro**
Espec. Liliana María Maturana Patarroyo***

Este artículo presenta los resultados de una investigación sobre discursos y prácticas evaluativas, cuyo fin es cualificar a los profesores de lenguas extranjeras de dos universidades colombianas en la evaluación de los aprendizajes. Principalmente se encontró: convivencia confusa de las evaluaciones sumativa y formativa; imprecisa definición del constructo lingüístico expresado en habilidades; énfasis excesivo en el área organizativa de la lengua a expensas del área pragmática; demasiada importancia a aspectos no lingüísticos; poca discriminación entre lo que se valora y lo que se mide en decisiones de promoción, y escasa consideración de muchas de las cualidades de las tareas evaluativas. Además, la falta de estabilidad laboral de los profesores, de directrices claras, de consenso, y de registros formales y sistemáticos de los resultados de las evaluaciones, contribuye a la dudosa promoción de los estudiantes.

Palabras clave: evaluación en lenguas extranjeras, evaluación formativa, evaluación sumativa, competencia comunicativa, tareas evaluativas.

This article presents the results of a piece of research on foreign language teachers' assessment discourse and practices carried out in two Colombian universities, with the aim of qualifying such practices. The main findings include: the confusing coexistence of summative and formative assessment; the imprecise definition of the linguistic construct in terms of skills; excessive emphasis on the organizational—not the pragmatic—area of language knowledge; the importance of non-linguistic aspects in pass/fail decisions; the indiscriminate use of assessment and testing, and the scarce consideration of several assessment task qualities. Additionally, the lack of teachers' work stability, of clear assessment guidelines and agreement, and of formal and systematic recording of assessment results, contributes to teachers passing students without the required communicative competence.

Keywords: evaluation in foreign languages, assessment, testing, formative assessment, summative assessment, communicative competence, assessment tasks.

L' on présente dans cet article les résultats d' une recherche concernant les discours et les pratiques évaluatives, et ce afin de proposer une formation ax professeurs de langues étrangères de deux universités colombiennes sur l' évaluation des apprentissages. L' on a pu constater principalement : une

* Recibido: 04-06-05/ Aceptado: 26-07-05

1 Esta investigación se realizó entre febrero del 2003 y agosto del 2004, gracias al apoyo financiero de la Vicerrectoría de Investigaciones y la Escuela de Idiomas de la Universidad de Antioquia, de la Facultad de Ciencias Humanas y Económicas de la Universidad Nacional de Colombia, Sede Medellín, y de la Vicerrectoría Académica, Departamento de Idiomas, de la Fundación Universitaria Luis Amigó.



Clara Inés Arias Toro • Liliana María Maturana Patarroyo

cohabitation confuse des évaluations sommatives et formatives, une définition imprécise du construct linguistique exprimé en habiletés langagières, une importance excessive donnée au champ organisatif de la langue au dépend du champ pragmatique, trop d' accent mis sur des aspects non linguistiques et une faible discrimination de ce qui est valorisé et ce qui est mesuré en termes de décisions de passage à un niveau supérieur, et aussi une faible considération d' un bon nombre de qualités des tâches d'évaluation. Le manque de stabilité d'emploi des enseignants, d'instructions claires, de consensus et d'enregistrements formels et systématiques des résultats d'évaluation contribue à un passage à un niveau supérieur quelque peu douteux.

Mots clés: évaluation de langues étrangères, évaluation formative, évaluation sommative, compétence communicative, tâches d'évaluation.



INTRODUCCIÓN

En el sistema educativo en general y en el universitario en particular, la evaluación de los aprendizajes es un tema que no se ha abordado con la profundidad que se requiere. Los programas académicos cambian siguiendo directrices y movimientos vanguardistas, pero la evaluación siempre queda rezagada. El momento actual exige, de los profesionales, el manejo de una competencia comunicativa que responda a requerimientos internacionales y locales para contratación laboral, posgrados, becas, pasantías, entre otros.

Esta investigación se centra en los discursos y las prácticas que sustentan el proceso de evaluación que utilizan los profesores, y en los factores que intervienen en su conceptualización, realización y en el desempeño de los estudiantes en las pruebas. La información que arroje esta primera fase del estudio (Arias *et al.*, 2003) servirá para avanzar en el diálogo colegiado y en la capacidad de propuesta y gestión, al elaborar con los profesores participantes un sistema de evaluación de los aprendizajes en lenguas extranjeras (LEs)² para los programas estudiados.

Richards (2005) afirma que los niveles alcanzados por estudiantes de LEs dejan mucho que desear en el contexto mundial y se han convertido en preocupación de muchos gobiernos. En nuestra ciudad y en el país, se reitera la falta de preparación de los estudiantes para desarrollar los objetivos de los cursos y se cuestiona el sistema de evaluación empleado, pues éste, la administración y los profesores parecen estar promoviendo estudiantes sin criterios claros y compartidos.

En este artículo planteamos primero las especificaciones de la investigación. En segundo lugar, presentamos el marco teórico; luego, los hallazgos seguidos de la discusión. Por último, planteamos conclusiones y sugerencias de este estudio. El marco teórico, los hallazgos y la discusión exponen las cinco categorías principales

2 Lengua extranjera: la lengua que se estudia o enseña en una comunidad donde no se emplea para la comunicación habitual y cotidiana entre sus usuarios. Segunda lengua: la lengua que se estudia o enseña en la comunidad donde es la lengua materna y, por ende, se emplea para la comunicación habitual y cotidiana (Richards, 2005; Dubin y Olshtain, 1986).



(evaluación formativa, evaluación sumativa, competencias, tareas evaluativas y cualidades) desde las perspectivas teórica, descriptiva y analítica, respectivamente.

EL ESTUDIO

Las preguntas de investigación que guiaron este estudio fueron: ¿cuáles son las prácticas evaluativas que se dan en los procesos de enseñanza y aprendizaje de inglés y francés como LEs?, ¿qué discursos sustentan las prácticas evaluativas que emplean los profesores en sus cursos?, ¿cuáles factores intervienen en el proceso de evaluación de los aprendizajes en dichos programas? y ¿cómo articular el conocimiento y la experiencia de los docentes con concepciones contemporáneas para lograr prácticas evaluativas consecuentes con el conocimiento de LEs que exige el momento actual? El objetivo general consistió en identificar y caracterizar discursos y prácticas evaluativas de los profesores de inglés y francés en los dos programas estudiados, con el fin de construir con ellos un sistema de evaluación que respondiera a las necesidades y exigencias de los procesos de enseñanza y aprendizaje de los citados idiomas en estos contextos.

El método de esta investigación educativa se enmarcó en los paradigmas interpretativo y sociocrítico (Arnal *et al.*, 1994) por varias razones. Para empezar, debido a que las prácticas y discursos evaluativos surgen de las interpretaciones que del tema tengan los involucrados, fue necesario explorar las subjetividades para que salieran a la luz el conocimiento, los valores, las creencias y las contradicciones de los profesores y los estudiantes (Marshall y Rossman, 1995). Segundo, las evaluaciones son decisorias en las vidas de las personas y las instituciones; por tanto, es necesario develar el poder que encarnan, para que el proceso de evaluación sea más justo, transparente y formativo (Shohamy, 2001). Tercero, la descripción de las prácticas evaluativas y los discursos de los profesores mostró los aspectos de evaluación, que requieren estudio y discusión. Esta información sentó las bases para el diseño de talleres de profundización conceptual y de capacitación en el análisis y diseño de la evaluación, que contribuirían al empoderamiento de profesores e investigadoras quienes, a su vez, desarrollarían capacidad de propuesta y gestión relacionada con el mejoramiento de la evaluación de los aprendizajes. Para lograr este objetivo profesional y social, fue esencial la participación de profesores y administradores en momentos clave del proceso.



La investigación se llevó a cabo en dos programas de dos universidades públicas. En una de ellas, los cursos de idiomas extranjeros no son parte de los programas académicos, sino de la formación integral. Sólo los estudiantes con el promedio académico más alto pueden tomar los cursos, pero todos deben pasar un examen de comprensión lectora como requisito de graduación del pregrado. En la otra universidad es necesario aprobar cuatro semestres de una lengua extranjera para obtener el diploma de pregrado. Ninguno de los programas tiene costo, pero cuando se reprueba un curso, el estudiante debe pagar si lo repite.

Los profesores de estos programas están todos contratados bajo la modalidad de cátedra o prestación de servicios y, por la falta de estabilidad laboral, deben trabajar en varias instituciones. La mayoría tiene cuatro y cinco grupos, algunos siete o más, lo que significa que cada profesor debe atender entre 60 y 180 estudiantes. La mayoría de los profesores son licenciados en idiomas o profesionales en áreas afines y han enseñado principalmente cursos básicos e intermedios a adultos. La mitad de ellos ha asistido a seminarios y conferencias sobre evaluación. Su edad oscila entre los 30 y 40 años.

Con el convencimiento de la importancia que tiene la implicación grupal en las decisiones relacionadas con el mejoramiento de las prácticas evaluativas y con la elaboración de un sistema de evaluación para los programas, el grupo de investigación invitó a profesores y administradores a reuniones informativas; veintiún profesores firmaron el código de ética propuesto por las investigadoras y se comprometieron por escrito a participar con: el suministro de instrumentos de evaluación que emplean (tradicionales: *quizzes*, exámenes, talleres; alternativos: formatos de autoevaluación y coevaluación; guías para evaluación de entrevistas) o asistencia a talleres; también con entrevistas semiestructuradas a ellos y a alguno de sus grupos. Los cinco talleres, con intensidad de cuatro horas cada uno, sirvieron para que los profesores se actualizaran en: tipos de evaluación y sus finalidades, cualidades de las pruebas, y áreas del lenguaje. A los talleres asistieron inicialmente veintiún profesores y administradores, y luego un promedio de doce. La disminución obedeció a motivos laborales, familiares, de viaje, o personales.

Para el análisis de datos seguimos las etapas propuestas por Burns (1999): a) recopilación de todos los datos, b) codificación y categorización, c) comparación,



d) interpretación y e) divulgación. La reunión de los datos incluyó la información proveniente de: pruebas e instrumentos de evaluación,³ entrevistas semiestructuradas individuales a diecisiete profesores, quince entrevistas semiestructuradas grupales a los cursos de los profesores entrevistados, y los lineamientos generales y programas de cursos. La categorización de todas las entrevistas surgió de los conceptos recurrentes en los datos y de la literatura sobre evaluación de los aprendizajes, categorías que se compararon con los instrumentos de evaluación entregados por los profesores. Las categorías o temas generales que surgieron fueron: evaluación formativa, evaluación sumativa, competencias, tareas evaluativas, y cualidades de las pruebas y tareas evaluativas. La divulgación inicial de los hallazgos e interpretaciones se hizo en un taller y esta confrontación contribuyó a la validez interna de la etapa de interpretación.

MARCO TEÓRICO

Tradicionalmente, la evaluación de los aprendizajes en LEs ha sido utilizada para recolectar información acerca del desempeño y de otros factores que influyen en el aprendizaje. Las prácticas evaluativas para recoger tal información varían en su propósito y para ello se cuenta con la **evaluación formativa y la sumativa**. Las tareas o procedimientos evaluativos empleados para allegar dicha información varían dependiendo de la definición que se tenga de la lengua y del qué, cómo, cuándo y para qué se va a evaluar.

La **evaluación formativa** se lleva a cabo durante el curso y sus resultados sirven para que el profesor ofrezca ayuda remedial oportuna o para que cambie el énfasis del curso, si fuera necesario. Dichos resultados también le sirven al estudiante para identificar sus áreas problemáticas (Alte Members, 1998). Los objetivos de la evaluación formativa son “la regulación pedagógica, la gestión de errores y el refuerzo de los éxitos” (Jorba y Casellas, 1997: 84). Lo anterior se logra a partir de una **evaluación continua** que incluye la **retroalimentación** y la **autoevaluación** (Hadji, 1997; Jorba y Casellas, 1997; Perrenoud, 1998; Instituto Cervantes, 2002). La

3 Frodden *et al.* (2004) presentan los hallazgos relacionados con las pruebas e instrumentos de evaluación suministrados por los profesores de los programas.



evaluación continua es un conjunto de actividades que brinda a los estudiantes la oportunidad de solucionar dificultades y buscar alternativas para demostrar y demostrarse lo aprendido. Aunque se realizan frecuentemente, hay discrepancia entre autores, pues según Landshère, citado en Meyer, 1995: “no se traduce(n) en una nota y menos aún en un puntaje” (21) y según Lafrancesco (1996), deben tener un valor porcentual o ponderado que será promediado al final.

La retroalimentación consiste en los comentarios orales o escritos que el profesor brinda al estudiante sobre su proceso de aprendizaje de la lengua. Aquella puede ser de carácter cognitivo, afectivo o ambos, y se realiza con el propósito de efectuar cambios o de redireccionar la enseñanza y el aprendizaje; debe ser relevante, completa y significativa, al lograr singularizar al estudiante y al tema tratado para que éste adquiera sentido e importancia a los ojos de aquel. Así mismo, los aspectos cognitivos o afectivos deben enmarcarse de tal manera que el estudiante y el profesor encuentren relaciones y comprendan la situación en su totalidad.

La autoevaluación es el proceso mediante el cual el alumno realiza una introspección crítica sobre sus debilidades y fortalezas en el proceso de aprendizaje; en el caso de las LEs, el alumno realiza valoraciones de su dominio lingüístico (Instituto Cervantes, 2002) y propone correctivos para su mejoramiento. La autoevaluación requiere del desarrollo de estrategias de metacognición que, unidas al autocontrol reflexivo de las acciones del sujeto que aprende, favorecen el aprendizaje (Hadji, 1997). Este tipo de evaluación también beneficia la enseñanza, pues los profesores la regulan, introduciendo cambios a partir de los progresos y problemas de aprendizaje observados en sus alumnos (Scriven, 1967, en: Jorba y Casellas, 1997). Por su parte, el estudiante aprende a autorregular su aprendizaje (Hadji, 1997), “aprendiendo a aprender” y volviéndose más autónomo (Benson, 2001) a partir del empleo continuo de la retroalimentación, la autoevaluación y la interacción social en el aula.

La evaluación sumativa se caracteriza por el momento en que se realiza y por su propósito. Se lleva a cabo después de la instrucción y las oportunidades de aprendizaje, para verificar el alcance de las adquisiciones logradas (Hadji, 1997); por tanto, es de carácter decisorio al definir qué tanto y qué aprendieron los estudiantes, y la promoción o no al final de un ciclo o período. En la mayoría de los casos,



la evaluación sumativa se basa en los resultados del conocimiento lingüístico de la lengua extranjera, no en los procesos. Tanto la evaluación sumativa como la formativa pueden hacerse con todo tipo de pruebas o tareas, tradicionales o alternativas, y su objetivo principal es dar cuenta confiable y válida de la competencia comunicativa.

En el área de las LEs, además de las competencias lingüísticas propiamente dichas, se han estudiado diversas **competencias generales** que se desarrollan durante toda la vida y afectan el desempeño lingüístico de los estudiantes. Ellas se muestran en la tabla 1.

TABLA 1. Comparación de competencias, características y aspectos —diferentes a la habilidad lingüística o a las competencias comunicativas— que afectan el desempeño de los estudiantes en pruebas de lengua

Marco de Referencia	Bachman y Palmer	Purpura
Competencias generales	Características de usuarios de la lengua	Factores que influyen los resultados en pruebas
Saber (Conocimiento): — Del mundo — Sociocultural: relaciones personales, valores, creencias, actitudes, convenciones sociales. — Conciencia intercultural.	Conocimiento temático: — Conocimiento del mundo. — Conocimiento cultural. — Conocimiento profesional, técnico o especializado. Características personales: — Edad, género. — Lengua materna, nacionalidad. — Condición o estatus de residencia. — Nivel educativo. — Preparación / experiencia con un tipo de prueba.	Factores demográficos: — Edad, género. — Lengua materna, etnia, nacionalidad. — Nivel socioeconómico y sociocultural. — Nivel educativo.
Saber hacer: — Destrezas sociales, de la vida, profesionales y de ocio. — Destrezas interculturales.	Competencia estratégica.	Estrategias metacognitivas.



Saber hacer: — Destrezas sociales, de la vida, profesionales y de ocio. — Destrezas interculturales.	Competencia estratégica.	Estrategias metacognitivas.
Saber ser: — Actitudes, valores, motivaciones. — Factores de la personalidad: locuacidad, parquedad; introversión, extraversión; temores; inteligencia; seguridad en sí mismo; autoestima; rigidez, flexibilidad.	Esquemas afectivos: aspectos afectivos o emocionales relacionados con el conocimiento temático. Características personales: — Aptitud para la LE.	Factores sociopsicológicos o socioculturales: — Actitud — Motivación — Esfuerzo. Factores relacionados con la personalidad: — Autoestima. — Ansiedad. — Capacidad para arriesgar. Factores cognitivos: — Aptitud.
Saber aprender: — Reflexión sobre la lengua y la comunicación. — Destrezas de estudio. — Destrezas heurísticas.	Competencia estratégica: — Definición de metas. — Evaluación o monitoreo. — Planeación. — Características personales — Estilo cognitivo.	Factores cognitivos: — Estrategias de aprendizaje o metacognitivas. — Estilo de aprendizaje: introversión, extraversión.

Fuente: Bachman y Palmer (1996); Purpura (1998); Instituto Cervantes (2002).

La **competencia comunicativa** (*communicative competence*), la define el *Multilingual Glossary of Language Testing Terms* como “la capacidad de utilizar adecuadamente el lenguaje en diversas situaciones y entornos” (Alte Members, 1998: 175).⁴ Bachman y Palmer (1996) se refieren al uso de la lengua y al desempeño en pruebas de lengua; dicho uso

[...] se enfoca, por un lado, en las interacciones entre las áreas de la habilidad lingüística, *language ability*, (conocimiento de la lengua y competencia estratégica), el conocimiento

4 Las traducciones del inglés o el francés son responsabilidad de las investigadoras.



temático y el componente afectivo y, por otro, en la manera como todos éstos interactúan con las características de la situación de uso de la lengua o prueba (62).

Estos autores y otros muy reconocidos (Canale y Swain, 1980; Widdowson, 1983; Instituto Cervantes, 2002) asumen la lengua y la capacidad para utilizarla como un conjunto de componentes que interactúan.

Según Bachman y Palmer (1996), el conocimiento lingüístico o el conocimiento del idioma comprende dos áreas: el conocimiento organizativo y el pragmático. El organizativo, referido a la manera como se organizan los textos orales o escritos, incluye el conocimiento gramatical y el textual. El conocimiento pragmático se refiere a la manera como las elocuciones, oraciones y textos se relacionan con los propósitos comunicativos del usuario de la lengua y a las características de la situación de uso de la lengua. Este conocimiento incluye el conocimiento sociolingüístico y el funcional.

Desde la década del sesenta, la **competencia comunicativa** se ha definido a partir de las habilidades lingüísticas, teniendo en cuenta el canal (audio, visual) y la modalidad (receptiva, productiva) (Lado, 1961, y Carroll, 1961, 1968, en: Bachman y Palmer, 1996). Esta concepción de competencia comunicativa distingue entre las habilidades lingüísticas (escucha, habla, lectura, escritura) y los componentes del conocimiento del lenguaje (gramática, vocabulario, fonología, grafología). Con el tiempo, la generalización en el empleo de estos términos ha llevado a la presunción de que se comparte su significado y lo que cada uno abarca, así como al reduccionismo de conceptos tan complejos. Bachman y Palmer (1996) y Bachman (1990) critican la concepción de competencia comunicativa basada en las cuatro habilidades, porque bajo una misma habilidad se incluyen actividades muy diferentes — como un *chat* y la lectura del periódico—, y porque se acostumbra practicar y evaluar las habilidades sin contextualizarlas, es decir, sin definir el componente pragmático.

Para evaluar la habilidad lingüística (*language ability*), estos autores proponen **tareas evaluativas** con características, áreas del conocimiento lingüístico y conocimiento temático bien delimitados.



Dichas tareas, según Bachman y Palmer (1996), pueden ser del dominio de la vida real, fuera de la situación de evaluación, o del dominio de la instrucción, cuando el uso de la lengua tiene propósitos de enseñanza, aprendizaje o evaluación como diarios, listas de cotejo, autoevaluación, portafolio, proyectos, talleres, presentaciones, entrevistas, juegos de roles y debates. Estos procedimientos alternativos, a diferencia de los tradicionales, requieren habilidades cognitivas más complejas y mayor producción por parte de los estudiantes; se enfocan en procesos y productos; son más flexibles y dinámicos; se aproximan a situaciones de la vida real; no interrumpen demasiado los procesos de enseñanza y aprendizaje al ser menos amenazantes y formales, e incluyen diversas competencias o habilidades (Brown, 1998; Brown y Hudson, 1998). Carroll (1993) llama estos procedimientos *soft* y Bustos (1997) se refiere a su carácter continuo que prioriza el proceso de aprendizaje. Los procedimientos tradicionales de evaluación, por su lado, son pruebas de papel y lápiz como exámenes, *quizzes* y talleres. Estos procedimientos, que Carroll (1993) denomina *hard*, se distinguen por su objetividad, confiabilidad, precisión, la primacía del producto sobre el proceso y porque los resultados se dan generalmente en términos cuantitativos.

Brown y Hudson (1998), quienes prefieren no distinguir entre formas de evaluación tradicionales y alternativas, optan por el término 'alternativas en evaluación', pues en cualquier caso se debe cumplir con las cualidades de evaluación responsable. Las tareas evaluativas o pruebas se constituyen en referentes prácticos o demostraciones tangibles del desempeño lingüístico de los estudiantes (así como de otros aspectos no lingüísticos), desempeño que se verá afectado también por las características de los mismos procedimientos.

Lo que se evalúa en las tareas (tradicionales o alternativas) se denomina **constructo** y la concepción que se tenga de la lengua y de la habilidad para comunicarse en ella, afecta su definición. El conocimiento de la lengua se considera como el **constructo lingüístico** (Bachman y Palmer, 1996) y los aspectos sociales, culturales, psicológicos y otros que afectan el desempeño lingüístico, podrían considerarse como **constructo no lingüístico**. Al elaborar pruebas o tareas evaluativas, los profesores deben definir el constructo, es decir, el conocimiento que el estudiante debe demostrar al realizarlas, ya que a partir de los resultados, harán inferencias relacionadas con el desempeño de los estudiantes en situaciones similares de uso de la lengua. La



definición precisa del tema, el propósito y el contexto de una prueba o tarea evaluativa será, entonces, el constructo que se va a evaluar.

Bachman y Palmer (1996) establecen seis **cualidades** para las pruebas tradicionales. La confiabilidad o fiabilidad, se refiere a la consistencia en la medición, bien sea entre jueces diferentes o en presentaciones diferentes de la prueba. La validez de constructo es la justificación de la interpretación del puntaje de un examen o del desempeño en una actividad evaluativa, o sea, la pertinencia de lo que se evalúa. La autenticidad se entiende como la relación entre las características de las tareas del examen y el uso de la lengua en la vida real, y la interactividad, como la intervención de la habilidad lingüística, el conocimiento temático y los esquemas afectivos del estudiante en las tareas de la prueba. El impacto se define como la influencia de los resultados de las pruebas en los ámbitos social, educativo y personal; por ejemplo, el efecto de arrastre (*washback effect*) en la enseñanza y el aprendizaje. Por último, la viabilidad se concibe como la relación entre los recursos disponibles y los recursos necesarios para el diseño y el uso de una prueba.

A la par de las anteriores cualidades, la introducción de prácticas evaluativas diferentes a las tradicionales es un primer intento para hacer del proceso de evaluación algo más democrático y justo, desde el diseño, hasta la administración y la interpretación de los resultados. Dicha introducción posibilita la evaluación y consideración de otros aspectos relacionados con el aprendizaje, los constructos no lingüísticos. Shohamy (2001) propone diversidad de fuentes, evaluadores y actores, y la negociación para un enfoque democrático, flexible y abierto, debido a las limitaciones de las pruebas tradicionales.

Shohamy (2001) y Bustos (1997) proponen otros niveles de rigor que deben cumplir las pruebas diferentes a las tradicionales: la justicia que asegura la legalidad, la ética y el respeto por el bienestar de los evaluados, la validez sistémica que guarda relación con la propuesta pedagógica, y la relevancia para que las pruebas reflejen el desempeño del estudiante en situaciones cotidianas y prácticas (no sólo en situaciones de papel y lápiz).

En conclusión, la tradición y los desarrollos en la evaluación de los aprendizajes en LEs ofrecen a los profesores múltiples perspectivas evaluativas, to-



das ellas con su imperativo de calidad marcado por las cualidades. Ya sea que las prácticas evaluativas se inscriban en lo formativo o sumativo, en lo tradicional o alternativo, o en combinaciones de éstas, la definición de las áreas del conocimiento lingüístico a evaluarse y la consideración de los aspectos que intervienen en los desempeños de los estudiantes son temas ineludibles, siempre que se busque la “reunión sistemática de evidencias para determinar si se producen los cambios esperados en cada alumno o no” (Tejada, 1999).

HALLAZGOS

Yo trato de evaluar todo (Prof. 18)

[...] evalúan todas las competencias que uno debe aprender del idioma [...] escucha, gramática, pronunciación (Est. Prof. 7).

[...] en la parte oral hemos evaluado diálogos, *roleplaying*, [...], presentaciones cortas [...] tengo una cita con los estudiantes [...] ellos deben haber leído un tema [...] que se hayan preparado y hayan buscado el vocabulario y que lo hayan entendido y que a través de sus propias palabras me hagan una descripción de lo que entendieron [...]; posteriormente le den una explicación al resto de la clase. [...] Eso me ha servido para evaluar muchas cosas: por una parte, la individualidad de los intereses y los gustos; y por otra, chequear cómo están las habilidades de escritura, de comprensión de lectura, ver los errores gramaticales y, ya en la parte relacionada con la presentación ante la clase, veo cómo está su fluidez, su pronunciación (Prof. 14).

Las citas anteriores reflejan la diversa y variada gama de propósitos, fuentes y herramientas existentes en la evaluación en LEs. Los hallazgos que surgen de esta investigación y su posterior discusión pretenden contribuir a aclarar y dilucidar algunos aspectos de la complejidad de esta evaluación.

‘Evaluar todo, las competencias, la preparación, los intereses, los gustos [...]’, son ejemplos de los propósitos que los profesores participantes manifiestan para la evaluación formativa, sumativa o ambas. Por un lado, si nos referimos a los programas, observamos que ninguno emplea el término evaluación formativa en los documentos disponibles para profesores y estudiantes, y sólo uno menciona la evaluación cualitativa de manera explícita. Sin embargo, ambos promueven la continuidad



en la evaluación, la retroalimentación y la autoevaluación, condiciones *sine qua non* de la **evaluación formativa**.

Los profesores, por su parte, también dan cabida en gran medida a la **evaluación formativa**; de los 17 profesores entrevistados, 14 se refieren a la evaluación continua; 11, a la retroalimentación; y 13, a la autoevaluación. Si examinamos la retroalimentación, los datos expresan la imaginaria de los profesores acerca de ella: 'apoyo psicológico', 'análisis de experiencias', 'percepciones', 'análisis del error en el papel del aprendizaje', 'toma de decisiones', 'recomendaciones', 'confrontaciones', 'es el momento de análisis de experiencias', 'sirve para crear conciencia, para fortalecer el compromiso', 'la oportunidad de analizar el papel del error en el proceso de aprendizaje'. Igualmente, los discursos muestran que existe un acuerdo tácito con relación al papel de la retroalimentación, al devenir en el medio real y concreto que vehiculiza la evaluación formativa, cumpliendo con las funciones de 'recomendar', 'sugerir', 'orientar' y 'exhortar' en aspectos afectivos y lingüísticos. Entre las modalidades de retroalimentación más usadas por los profesores están la entrevista, la asesoría, la entrevista individual, la puesta en común, la conferencia y las observaciones.

Los profesores caracterizan la retroalimentación como 'una confrontación' entre estudiante y profesor, y una 'construcción permanente' del proceso de aprendizaje. Este carácter de permanencia da paso a otro componente que caracteriza la evaluación formativa, la continuidad. "Opté por trabajar con ellos todo el rato haciendo seguimiento [...] que incluye [...] la participación en clase, la asistencia [...] los valores y, en la parte cognitiva, [...] ponía un taller o [...] una evaluación corta" (Prof. 8).

La autoevaluación —que recoge aspectos lingüísticos y no lingüísticos—, se considera como un momento para analizar, reflexionar y expresar 'qué hizo bien y qué no', 'sus fortalezas y debilidades', 'participación', 'interés', 'uso del tiempo', en otras palabras, para fomentar el saber aprender, la competencia estratégica o las estrategias de aprendizaje. Estas estrategias representan 'el trabajo extraclase' y el 'cómo ve el estudiante su competencia comunicativa'.



La **evaluación sumativa** aparece en los programas en momentos claros para verificar los aprendizajes. Esta medición o evaluación sumativa —cuantitativa o cualitativa—, está expresada en la realización de dos exámenes por nivel, el parcial y el final, con parte oral y escrita.

La evaluación es cualitativa, es un proceso continuo que se realiza a través de las diferentes actividades. Se evalúa en cada actividad el proceso del estudiante y se da una retroalimentación del proceso. Al finalizar el semestre el alumno recibe un informe final de su desempeño durante el nivel, y de acuerdo a éste será promovido o no al siguiente nivel (Programa de curso).

Los profesores, sin mencionar el término evaluación sumativa, dicen que hay momentos de corte en los cuales evalúan, 'cada cuatro unidades' y 'el examen parcial y el final' (Profs. 1, 8, 20), con el fin de revisar los logros hasta ese punto. La **evaluación sumativa**, por definición, se distingue por la función que cumple en un momento preciso, no tanto por los términos que se emplean para reportar los resultados. Los términos cuantitativos (3,0 / 5,0, 7 / 10, etc.) como los cualitativos (excelente, regular, deficiente o aprobó, reprobó) y la combinación de éstos, expresan la decisión de promover o no al estudiante a algún nivel superior. El corte puede hacerse en puntos intermedios o al final; nuevamente, algunos profesores dicen dar una nota o concepto a constructos lingüísticos 'gramática y vocabulario' y a constructos no lingüísticos, como 'asistencia', 'interés' y al trabajo que sobre el aprendizaje realizan los estudiantes, de lo cual éstos también dan testimonio.

En el otro programa, la evaluación está definida por componentes (clase taller, asesorías, centro de recursos, laboratorio de lectura, proyecto) y la administración ha definido los porcentajes para cada uno, algunos con base en la asistencia (ej., asesorías, centro de recursos) y otros, en los resultados (ej., clase taller, proyecto, laboratorio de lectura). El desempeño se expresa en conceptos y la aprobación del curso depende del mayor número de los logros alcanzados (80%) con criterios de excelente, bueno y aceptable.

A la par con estos tipos de prácticas, aparece en el discurso de algunos docentes 'lo cualitativo' y 'lo cuantitativo'; el uso de 'número', 'nota', 'porcentajes', 'fraccionarios', 'valores', 'rango', 'equivalente', demuestra la presencia de una visión cuanti-



tativa respecto al desempeño de los estudiantes. El uso de 'indicadores de logro', 'anotaciones', 'criterios' acusa un manejo más cualitativo de la evaluación. Para varios profesores, la combinación de un descriptor cualitativo con un equivalente numérico sustenta de cierto modo la práctica evaluativa; en este sentido, se encuentran los 'descriptores definidos en términos de frecuencia', 'nota cuantitativa en quiz', 'equivalente a (nota) cualitativa para el alumno'. Para otros profesores, el no uso de la palabra 'número', parece traducirse automáticamente en evaluación cualitativa. En este sentido se encuentran 'rayas', 'puntos', 'símbolos', 'bueno', 'aceptable', 'excelente', 'insuficiente', 'aprobó', 'reprobó'. Así, en medio de esta diversidad y confusión terminológica, conviven la evaluación formativa y la sumativa en la práctica de los profesores. "Yo evalué todo el tiempo cuantitativamente... y al final, aprobó o no aprobó" (Prof. 8). "Los puntos y rayitas definen lo malo o bueno" (Prof. 12).

De igual forma, evaluar 'la gramática', 'la comprensión de lectura', 'la fluidez', 'el interés'... representa el **constructo** evaluado por los profesores. Cuando a éstos se les solicitó explicitar los aspectos lingüísticos que evaluaban, las respuestas más precisas se referían a los contenidos y objetivos de los cursos, gramática, vocabulario y conocimiento temático, pronunciación, coherencia y cohesión (todos componentes del área de conocimiento organizativo, según Bachman y Palmer, 1996), a 'argumentar un punto de vista' (componente funcional, Bachman y Palmer, 1996) y al 'componente cultural' (lo pragmático y sociolingüístico, Bachman y Palmer, 1996). No emplearon términos lingüísticos precisos, solamente 'las cuatro habilidades', 'comprensión de lectura y habilidad oral', entre otros.

En lo que evalúan o se evalúa, los profesores, sus estudiantes y las pruebas suministradas presentan gran variedad de aspectos afectivos, socioculturales y sociopsicológicos o competencias generales (Bachman y Palmer, 1996; Purpura, 1998; Instituto Cervantes, 2002). Profesores y estudiantes se refieren a estos aspectos con términos émicos como: 'asistencia', 'creatividad', 'iniciativa', 'recursividad', 'cumplimiento', 'esfuerzo', 'motivación', 'participación y estrategias de estudio'. Los lineamientos de los programas también fomentan abiertamente algunos de estos aspectos no lingüísticos: 'investigación', 'autoestudio', 'trabajo grupal', 'trabajo responsable', 'asistencia', 'estrategias para desarrollar (las) habilidades'.



Debido a que los profesores equipararon la competencia, las habilidades, las subhabilidades, la motivación y el interés al constructo, se hizo difícil hallar claras evidencias para cada uno de los conceptos. Por este motivo, el equipo decidió examinar las **tareas evaluativas** descritas en las entrevistas y evidenciadas en los instrumentos entregados por los profesores, ya que éstas son manifestaciones concretas de lo que se evalúa: el constructo.

Los profesores y los estudiantes mencionaron una amplia gama de procedimientos evaluativos que aplican continuamente y que van desde los tradicionales (talleres, dictados) hasta los alternativos (debates, entrevistas, proyectos), desde simples (talleres) hasta complejos (proyectos). Algunas de las tareas alternativas son similares a las empleadas para enseñar y practicar la lengua extranjera, como juegos, canciones, diálogos, dictados. Hay tareas con instrumentos como lista de cotejo, guía para un video, y sin instrumentos, como los diálogos; así como con criterios previamente establecidos y comunicados, y sin criterios.

La finalidad evaluativa de las tareas también es diversa: se emplean para autoevaluación y retroalimentación (evaluación formativa) o heteroevaluación y promoción (evaluación sumativa), quedando la duda al respecto. En un programa, las tareas evaluativas están institucionalizadas (asesoría, proyecto, etc.) y tienen un porcentaje establecido para la aprobación del curso; en el otro no hay porcentaje definido, y se evalúan generalmente de manera cualitativa.

Los profesores y los estudiantes mencionan con mucha frecuencia los proyectos y sus respectivas presentaciones (exposiciones). El proyecto se compone de múltiples episodios: “[...] proyecto [...] eligen el tema de la carrera que estudian, [...] los estudiantes investigan, [...] traen informes [...] cronograma de investigación” (Prof. 21).

Con muy baja presencia aparecen trabajos, cartas, composiciones, diálogos (grabados o no), juegos de roles, dictados, debates y transcripciones fonéticas.

En muy pocas ocasiones se mencionan formatos de evaluación que acompañen las tareas evaluativas: las tareas *soft*, en su mayoría, no cuentan con instrucciones claras ni criterios definidos. Tampoco abundan instrumentos tangibles con instrucciones, criterios y puntuación que acompañen y antecedan a las tareas evaluativas, ni



registros con convenciones explícitas y compartidas. Los sistemas idiosincrásicos empleados por varios profesores dan cuenta de ello: “llevo un registro mental” (Prof. 3) y “memorizo la participación de cada uno, si contestan a las preguntas” (Prof. 4).

Por último, ‘búsqueda de vocabulario’, ‘la preparación’, etc., reflejan la intención implícita del profesor de promover **competencias estratégicas** en el estudiante y, a su vez, la intención expresa de evaluarlas. “También hay otros elementos que evaluamos [...] no solamente de la lengua como tal, sino todo lo demás que tiene que ver con [...] actitud, [...] o sea, mostré interés, le dediqué tiempo suficiente [...]” (Prof. 5). “Una nota final o un concepto final tiene que ver con el esfuerzo que cada quien hace sobre sus propias debilidades para aprender la lengua” (Est. Prof. 2).

Lo anterior quiere decir que al momento de promediar las notas o conceptos de los estudiantes, también se incluyen este tipo de aspectos. Las actitudes y valores, y la competencia estratégica, no han sido definidos con antelación ni en los criterios de promoción, ni como constructo en las pruebas, y aparecen como ‘nota de concepto’ o ‘nota de participación’ que el profesor define al final como apreciación. El concepto ‘actitudes y valores’ escasamente se encuentra en las pruebas provistas por los profesores y sólo un profesor entregó fichas de autoevaluación que emplea como instrumentos. Entre los criterios que definen la promoción de un estudiante, muchos de los aspectos no lingüísticos intervienen en la promoción de manera arbitraria y subjetiva. Sin embargo, aunque la imagen que tiene el profesor frente a las actitudes, valores y estrategias no es del todo clara, sí se vislumbra en su discurso un reconocimiento explícito de que todo lo que afecte la acción educativa es susceptible y debe ser motivo de evaluación. Este fenómeno, que a simple vista parece contradictorio con lo promovido en la filosofía de los programas, se da cuando los profesores no establecen ni explicitan que los aspectos no lingüísticos también hacen parte de su constructo evaluativo y, por consiguiente, deberían recibir el mismo tratamiento que se le da al constructo lingüístico.

DISCUSIÓN

De algunos instrumentos de evaluación entregados por los profesores y de las entrevistas a profesores y estudiantes, la **evaluación formativa** parece tener un sitio en ambos programas. Si bien la coevaluación (5 de los 17 profesores entrevista-



dos) no se favorece tanto como la **autoevaluación** y la **retroalimentación**, profesores y estudiantes ven en dichas prácticas la oportunidad para recibir comentarios, reflexionar e introducir cambios en su accionar.

La autoevaluación, casi siempre propuesta por el profesor, aunque cumpla en cierta medida con lo que se requiere para que el estudiante se autorregule y favorezca su aprendizaje —estrategias de metacognición y autocontrol reflexivo— (Hadji, 1977), no parece haber redundado en aprendizajes mejorados. Aunque la autoevaluación y la retroalimentación están presentes en la mayoría de los discursos de los profesores participantes, no se ven evidencias de haber contribuido a la autonomía y al aprender a aprender de los estudiantes. Esta carencia fue precisamente la razón que impulsó esta investigación, a saber, el paso frecuente de estudiantes a niveles superiores sin tener los conocimientos requeridos. La presencia de la autocalificación, la cual dista mucho de la autoevaluación y su propósito, puede explicar en parte esta problemática.

La **continuidad**, otro elemento clave en la evaluación formativa, también presente en el discurso de los profesores y los estudiantes, deja mucho que desear, al no estar siempre acompañada de registros formales con especificaciones claras de lo que se va a evaluar y de la manera como se va a evaluar.

¿Dónde radica la falla de este proceso formativo? A partir de los testimonios y de algunos de los instrumentos entregados por los profesores, parece que cada uno de los objetivos que persigue la evaluación formativa sí se alcanzara en cierta medida. Hay indicios de regulación pedagógica, de gestión de errores y de refuerzo de fortalezas (Jorba y Casellas, 1997), pero no parece haber conexión entre estas acciones. La razón puede ser que la continuidad, la retroalimentación y la autoevaluación, aunque se presentan en el proceso, pueden no estar planeadas de manera coherente ni relacionadas con propósitos establecidos, afectando los resultados y alcances de estas prácticas. Algunas de las interpretaciones que se podrían dar a tal situación son de tipo práctico: una, como profesores cátedra, la mayoría no cuenta con el tiempo suficiente para hacer la planeación y el seguimiento que este tipo de evaluación requiere. Otra, la falta de claridad institucional frente a este tipo de evaluación hace que, en muchas ocasiones, el profesor sólo la incluya como una obligación, no como una necesidad educativa.



Vemos que la evaluación formativa, más que tener como condición *sine qua non* la retroalimentación y la autoevaluación a partir de una continuidad, requiere de una planificación organizada, de una unificación de propósitos y de una sincronía entre las prácticas. Los profesores no hablan de evaluación formativa como tal, sino de algunos de sus elementos, los cuales parecen emplear de manera aislada y bastante al azar. Solamente uno de los profesores dice emplear los tres componentes y habla de 'seguimiento' (Prof. 5). Aun cuando lo más cercano es lo que denominan evaluación cualitativa, no se observa que los profesores dimensionen lo que implica lo formativo en su conjunto, sus requerimientos, su intención y su alcance.

Los profesores se refieren a la evaluación que lleva nota numérica como cuantitativa, y como cualitativa a la evaluación continua que se expresa con descriptores y brinda retroalimentación. Estas prácticas de la evaluación formativa no se constituyen en evaluación formativa como tal, porque los profesores no las emplean todas, no manifiestan realizarlas de manera sistemática y planificada, ni aprovechan todo el potencial que ofrecen; tampoco hay claridad y unidad de criterios entre los profesores de un mismo programa. No se puede negar que, en términos de la evaluación formativa, los profesores perciben su sentido de totalidad, pero no de unidad. Los diferentes componentes de este tipo de evaluación se tienen en cuenta de manera aislada, pero no se percibe intención de agruparlos con fines realmente pertinentes a la evaluación totalizante, integral y holística.

La convivencia de lo **sumativo** y lo **formativo**, de lo cualitativo y lo cuantitativo en las prácticas evaluativas de los profesores participantes, acusa, por un lado, un deseo integrador de diversas prácticas que impriman a la evaluación un carácter más alternativo, abierto y flexible; y por otro, la confusión existente en torno al uso y manejo de dichas prácticas. La evaluación cualitativa se lee como formativa y la sumativa se traduce en cuantitativa por el uso de números. Es claro que al final del curso, a lo cuantitativo o cualitativo que tengan del desempeño (lingüístico y no lingüístico) de los estudiantes, le aplican algún tipo de operación que da como resultado la evaluación sumativa, o sea, aprobó, no aprobó, siguiendo las instrucciones de los programas.



La inconsistencia entre los resultados reales y el proceso muestra un bache y una relación que parecería irreconciliable entre lo formativo y lo sumativo, lo cuantitativo y lo cualitativo. Si esta relación se considerara desde lo holístico, integral y dinámico, rompería de una vez por todas con la separación casi satanizada o estigmatizada entre lo cualitativo y lo cuantitativo, entre lo formativo y lo sumativo. Las prácticas evaluativas dinámicas y permanentes que aparecen como apoyo de todo el proceso, sufren un reduccionismo tenaz al resolverse simplemente con una operación matemática. Los resultados, en números o palabras, opacan todo el proceso.

Adicionalmente, para la decisión de promover o no a los alumnos, hay profesores que al emplear descriptores, logros u objetivos, tienen como punto de referencia la práctica generalizada de la 'mitad más uno' y, si no se cumple, en algunos casos completan con algún aspecto no lingüístico como el interés: "[para la promoción] que [el muchacho] cumpla esos objetivos, [...] le vi un buen interés, [...] pasémolo [sic] porque hay interés, entonces de diez cumplió cinco, vamos a ver qué pasa en el cuarto nivel" (Prof. 6). Este ejemplo demuestra el nivel de inconsistencia interna y externa entre lo que se valora (*assessment*), y lo que se mide (*testing*). Aquello que parece valorarse durante el proceso, adquiere el tinte de medición al momento de tomar decisiones.

Sucede que estos conceptos están tan arraigados en la educación, sin la debida claridad, que tratándose de números o de descriptores, profesores y estudiantes, inconscientemente, han desarrollado un sistema de equivalencias que termina por borrar la diferencia entre cantidad y calidad. Lo anterior se da, a pesar de la preferencia de muchos por lo cualitativo. La confusión en el tema de la evaluación la exacerbaban las directrices de estos dos programas, contradictorias y opacas, al fomentar la evaluación cualitativa y continua, dar porcentajes e incluir constructos lingüísticos muy amplios y constructos no lingüísticos de todo tipo. Aunque en la evaluación formativa los comportamientos tienen un papel importante, la confusión se vuelve más fuerte cuando no hay criterios para favorecerlos, cuando no se tiene claridad sobre su papel en la evaluación o cuando se reduce su conceptualización. En estos casos, la creatividad, la responsabilidad y la participación devienen 'ases bajo la manga' para inclinar la balanza final en momentos decisorios.



Los profesores de los programas estudiados emplean procedimientos de evaluación tradicionales o *hard*, así como otras alternativas en evaluación o procedimientos *soft*. En cualquier tipo de procedimiento evaluativo, la definición del constructo lingüístico que se va a evaluar se debe realizar a la par con la descripción de la **tarea** mediante la cual se evaluará. Por esto es que la especificación precisa de las características de las tareas (tema, propósito, contexto) se constituye en el constructo que se quiere evaluar. De esta manera, la tarea es el referente práctico que posibilita que el estudiante demuestre su competencia lingüística. Para varios profesores de este estudio, las tareas parecen ser el fin en sí mismas y no el medio para obtener la información para la evaluación. En general, a las tareas y actividades evaluativas que los profesores mencionan les falta pertinencia en el proceso de evaluación, coherencia, consenso con relación a lo que involucran y a su papel evaluativo, y una planificación que integre la descripción de las características sociolingüísticas y funcionales a las áreas del conocimiento lingüístico que se va a evaluar. El proyecto sí tiene etapas definidas y compartidas por la mayoría de los profesores; el problema, como en la mayoría de las tareas alternativas, radica en la inclusión de demasiados constructos, sin especificar cuáles entran a cumplir un papel decisivo en la evaluación. Tampoco se explicita el peso de cada uno de ellos, lo cual se complica con la tendencia a priorizar las competencias generales y los aspectos socioculturales y psicológicos, cognitivos y de la personalidad sobre el constructo lingüístico.

La similitud entre las tareas evaluativas y las de enseñanza que se encuentra en las entrevistas de los profesores puede tener su lado positivo y su lado negativo. Lo positivo es que las tareas evaluativas alternativas intimidan menos al estudiante que las tradicionales, porque se realizan en la cotidianidad sin interrumpir el flujo de las clases y pueden ser más auténticas y naturales (Bustos, 1997). Lo negativo tiene que ver con que, si no se anticipan el propósito y el constructo que se tendrán en cuenta, esas tareas evaluativas pierden validez. Igualmente, su viabilidad y fiabilidad se ven afectadas; la primera, por el tiempo y atención que requieren del profesor, y la segunda, porque hay un solo evaluador y los criterios pueden variar con mucha facilidad de una tarea a otra, de un día para otro o de un grupo a otro. Si, como lo expresan profesores y estudiantes, se está evaluando continuamente, unos y otros pueden confundirse y no diferenciar las situaciones de evaluación de las de enseñanza y aprendizaje, lo que parece reflejarse más



bien en una situación de evaluación informal, imprecisa y laxa, todo lo cual afecta la transparencia y el impacto de la tarea. Esta confusión se da en el campo de las tareas auténticas usadas para el aprendizaje y en las tareas auténticas usadas con fines evaluativos. El profesor con regularidad interpreta como tarea auténtica lo que proviene de un texto auténtico. Una receta de cocina, por ejemplo, es tomada como auténtica y se evalúa por el despliegue en la preparación y la pronunciación del estudiante, no por lo que sepa hacer con la lengua a partir de la tarea. El seguir instrucciones y utilizar conectores secuenciales son tareas de la vida real; si este fuera el constructo de la prueba, sería una tarea auténtica e incrementaría su validez de constructo.

Esta indefinición y falta de claridad se complica al unirse a la casi total ausencia de formatos, criterios y consensos entre el colectivo docente. Podría decirse que las tareas generalmente se evalúan de manera arbitraria pues, al no contar con formatos y registros con criterios de evaluación detallados, se incluyen aspectos lingüísticos y no lingüísticos al amañó del profesor. Así, la evaluación de los desempeños se presenta como global, subjetiva y falta de rigurosidad. Otras posibles fuentes de confusión son: una, que los profesores relacionan las tareas alternativas con la formación del estudiante y no necesariamente con la evaluación formativa para mejorar la lengua extranjera; y otra, que no parecen considerar las pruebas tradicionales como oportunidades para la formación, sino más bien para la promoción. Falta comprender que las tareas evaluativas alternativas y las tradicionales pueden servir tanto para la evaluación formativa como para la sumativa, pues lo que importa es que el propósito y el constructo estén claros.

Con relación al **constructo** —la piedra angular a evaluar en todo tipo de tareas, en las entrevistas—, se observó que los profesores desplegaron una amplísima gama de términos para referirse a los aspectos no lingüísticos, en tanto que el vocabulario para el constructo lingüístico se vio limitado a las cuatro habilidades y unos pocos componentes de la lengua, especialmente del área organizativa. Se da tanto una reducción como una generalización en lo atinente al constructo. La reducción se da porque se privilegia el área organizativa del conocimiento lingüístico (vocabulario, sintaxis, pronunciación, cohesión, organización textual); del área pragmática se evalúa la función ideacional, o sea, la que permite el intercambio de información en relaciones de transacción más que de trato interpersonal (Brown y Yule, 1983; Yule, 1996;



Richards, 2005). Este último componente del conocimiento pragmático sí se encuentra en los documentos empleados en los dos programas, lo cual nos lleva a preguntarnos acerca de su escasa presencia en los discursos estudiados y en la evaluación. Los componentes funcionales y sociolingüísticos —que se adquieren en la lengua materna— siempre están presentes cuando se emplea una lengua en una situación comunicativa (Brown y Yule, 1983; Yule, 1996). El profesor decide los aspectos pragmáticos a incluir, pero su poca mención puede obedecer a que se han vuelto inconscientes, a alguna incomodidad con ellos o a no saber nombrarlos. Esta ausencia puede explicarse también por la falta de programas de inducción y capacitación coherentes y continuos, y por el poco diálogo colegiado entre profesores de cátedra, quienes fueron los participantes en este estudio.

La generalización del constructo se manifiesta en el empleo de las cuatro habilidades y algunos componentes de la lengua. Lo significativo de la cuestión parece ser que, con el tiempo, el concepto 'habilidades' se ha constituido no sólo en un ejemplo de economía del idioma, sino en parte fundamental del discurso de profesores e investigadores, al punto que pareciera que se tratara de un imaginario colectivo, de una representación que parece compartirse en el campo de la enseñanza y el aprendizaje de las LEs. No obstante, a raíz de este estudio, nos han surgido dudas y cuestionamientos relacionados con el nivel al cual realmente se comparte esta representación, pues aparentemente se está de acuerdo pero, cuando se solicita más explicación, hay ambigüedad y falta de consenso.

La generalidad y la falta de precisión de 'lo que se evalúa' son elementos que complican la búsqueda de consensos y explicaciones en el tema de la evaluación en los programas estudiados. Adicionalmente, esta falta de precisión puede llegar a afectar la promoción, ya que se tornan cuestionables los resultados de las evaluaciones y los mismos criterios de promoción pues, lo que se concluye con relación a las cuatro habilidades lingüísticas y a los factores no lingüísticos de los estudiantes, no indica realmente lo que pueden hacer o no con la lengua.

A la mayoría de los procedimientos evaluativos *soft* les falta transparencia, pues con frecuencia carecen de instrumentos e instrucciones; cuando los hay, falta claridad en la puntuación, el peso de los componentes y los criterios. Los registros de esta evaluación son escasos e, igualmente, sin convenciones claras y compartidas,



como lo muestran los sistemas idiosincrásicos empleados por varios profesores (rayitas, bolitas, etc.).

Podemos afirmar, entonces, que las prácticas evaluativas de los profesores no son un corpus coherente, cohesionado, consciente y consensuado. Al analizar estas limitaciones y carencias entre profesores con un buen grado de preparación y con gran creatividad en el diseño de tareas evaluativas, arriesgamos varias explicaciones. Falta que los profesores potencien al máximo tales tareas, acompañándolas de la especificación de componentes de la lengua susceptibles de evaluarse de manera puntual. Los datos escasamente presentan esto —lo cual daría cuenta de una evaluación democrática, justa y transparente, así como confiable y válida—. La elaboración de instrumentos de evaluación requiere de tiempo y recursos, de lo cual carecen estos profesores de cátedra. Otra razón podría ser el desconocimiento de las características y cualidades requeridas en el diseño de instrumentos evaluativos. Si bien es cierto que, en general, los programas de Licenciatura del país incluyen contenidos sobre la evaluación de los aprendizajes y que la mitad de los profesores entrevistados han realizado estudios relacionados con la evaluación, puede que éstos no hayan sido suficientes, hayan sido demasiado teóricos, se hayan centrado en el diseño de las pruebas tradicionales de papel y lápiz, o no se haya conectado lo teórico con el diseño práctico.

CONCLUSIONES

Esta investigación nos permitió penetrar en la intimidad de las prácticas evaluativas de los profesores y las diversas concepciones, creencias y limitaciones que las condicionan. Este conocimiento debe contribuir a aceptar carencias, esclarecer creencias y empezar a asumir cambios necesarios para apostarle a una evaluación más adecuada a lo que el presente y el futuro requieren de los profesores y estudiantes de LEs.

Ante todo fue reiterativa la imprecisa definición del constructo lingüístico, la concentración en el área organizativa y el olvido del área pragmática del conocimiento lingüístico. La falta de definición de las características de las tareas lingüísticas y de sus propósitos de instrucción o de evaluación sumativa o formativa, de medición o valoración, ocasionó la inclusión de aspectos no lingüísticos que terminaron imponiéndose a la hora de tomar decisiones de promoción. Al mismo tiempo, la falta de



planeación, consenso y transparencia ocasionó dificultades en su diseño, aplicación y en la evaluación del desempeño.

La evaluación formativa no se distingue con claridad de la sumativa; por tanto, la retroalimentación y la autoevaluación no se realizan con el rigor, la sistematicidad y continuidad necesarios. En consecuencia, difícilmente redundan en el mejoramiento de los procesos de enseñanza y aprendizaje, su razón de ser.

Falta consenso entre los profesores y en las directrices de los programas en el tema de la evaluación. En general, se carece de registro formal y sistemático de los resultados de las evaluaciones, lo cual dificulta el desarrollo de los programas. El uso indiscriminado y arbitrario de lo que se valora (*assessment*) y de lo que se mide (*testing*) en momentos de corte y de decisión, contribuyó al avance dudoso de estudiantes a unidades o niveles superiores.

Los resultados de esta investigación son el insumo para la elaboración conjunta de un sistema de evaluación para los programas. No se alcanzó a diseñar dicho sistema, pues en el transcurso del proyecto observamos que en el tema de la evaluación de los aprendizajes en LEs hay mucho camino por recorrer en lo concerniente a un lenguaje común, a capacitación y actualización, y al compromiso decidido y efectivo de la administración para lograr continuidad y sostenibilidad en el proyecto. Sin embargo, esto que parece una debilidad, se convirtió en una fortaleza, porque se llevó a cabo un seminario-taller⁵ con algunos profesores participantes y profesores de otro programa de una de las instituciones ya involucrada en la investigación. Con este evento se conformó la comunidad de coinvestigadores para la segunda fase: dos profesores representantes de los programas de la primera fase, cuatro del nuevo programa y las coinvestigadoras de la primera fase. Los profesores de los programas implementarán el sistema, el equipo lo evaluará y posteriormente lo presentará como innovación curricular a las instituciones participantes.

En el equipo de investigación nos quedan muchos interrogantes relacionados con nuestro papel como evaluadores de LEs. ¿Hasta qué punto nos vemos los profesos-

5 Seminario-taller de veinte horas presenciales y veinte de trabajo independiente. Temas: tipos de evaluación, constructo, competencias, cualidades de las pruebas / tareas evaluativas, elaboración del sistema de evaluación.



res como evaluadores influenciados por la imagen que queremos proyectar de nosotros mismos ante la comunidad, y hasta qué punto por aquélla que no queremos proyectar? ¿Hasta dónde podrán la preparación y la claridad en el tema de la evaluación contribuir a disipar esos temores para realizar nuestro papel como evaluadores de manera justa y democrática? ¿Qué nivel de honestidad y de madurez necesitamos alcanzar para realmente nutrirnos con los resultados de nuestra labor y regular los ambientes de aprendizaje y autorregularnos como maestros?

Una de las limitaciones de esta investigación fue que no realizamos observaciones de clase; por ello, no podemos referirnos a la regulación de la enseñanza, ni a la autorregulación del aprendizaje. Escasamente encontramos estos procesos en sus discursos. En la segunda fase de esta investigación acogeremos diversas técnicas de recolección de datos para ampliar estos temas.

Por encima de todo, y para finalizar, el aprendizaje de la lengua extranjera por los estudiantes se constatará cuando se pueda verificar, valorar y certificar con el consenso y el rigor que éste merece.

BIBLIOGRAFÍA

- Arias, Clara Inés *et al.*, 2003, *Prácticas evaluativas en la adquisición de LEs*: Manuscrito del proyecto presentado al CODI, U. de A.
- Alte Members, 1998, *Multilingual Glossary of Language Testing Terms*, Cambridge, Cambridge University Press.
- Arnal, Justo *et al.*, 1994, *Investigación educativa: fundamentos y metodologías*, Barcelona, Labor.
- Bachman, Lyle F., 1990, *Fundamental Considerations in Language Testing*, Oxford, Oxford University Press.
- Bachman, Lyle F. and Adrian S. Palmer, 1996, *Language Testing in Practice- Designing and Developing Useful Tests*, Oxford, Oxford University Press.
- Benson, Phil, 2001, *Teaching and Researching Autonomy in Language Learning*, London, Pearson Education Limited.
- Brown, James D., 1998, *New Ways of Classroom Assessment*, Alexandria, Teachers of English to Speakers of Other Languages.
- Brown, James D. and Thom Hudson, 1998, "The Alternatives in Language Assessment", *TESOL Quarterly*, 32 (4), Bloomington, Illinois.



- Brown, James D. and George Yule, 1983, *Teaching the Spoken Language*, Cambridge, Cambridge University Press.
- Burns, Anne, 1999, *Collaborative Action Research for English Language Teachers*, Cambridge, Cambridge University Press.
- Bustos, Félix, 1997, *Mutaciones en evaluación y lineamientos del Ministerio: análisis y crítica*.
- Canale, Michael and Merrill Swain, 1980, "Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing", *Applied Linguistics*, 1 (1).
- Carroll, Brendan J., 1993, "Typologie des textes de langue", *Le Français dan le Monde, numéro spécial: Évaluation et certifications en langue étrangère*.
- Dubin, Fraida and Elite Olshtain, 1986, *Course Design: Developing Programs and Materials for Language Learning*, Cambridge, Cambridge University Press.
- Frodden, Cristina *et al.*, 2004, "Analysis of Assessment Instruments Used in Foreign Language Teaching", *Íkala, revista de lenguaje y cultura*, Medellín.
- Hadji, Charles, 1997, *L'évaluation démystifiée*, Paris, ESF Éditeur.
- Ianfraceso, Giovanni M., 1996, Nueve problemas de cara a la renovación educativa; alternativas de solución, Libros y Libres.
- Instituto Cervantes, 2002, *Marco de referencia europeo para el aprendizaje, la enseñanza y la evaluación de lenguas*, [documento electrónico], (traducción al español de Alejandro Valero Fernández), <<http://cvc.cervantes.es/obref/marco/>>, [consulta: 23 abr 2005].
- Jorba, Jaume y Ester Casellas, 1997, *La regulación y la autorregulación de los aprendizajes*, Madrid, Síntesis.
- Marshall, Catherine and Gretchen B. Rossman, 1995, *Designing Qualitative Research*, London, SAGE Publications.
- Meyer, G., 1995, *Évaluer: pourquoi? et comment?*, Paris, Hachette.
- Perrenoud, Philippe, 1998, *L'évaluation des élèves. De la fabrication de l'excellence à la régulation des apprentissages: entre deux logiques*, Paris, De Boeck & Larcier S.A.
- Purpura, James E., 1998, *Learner Strategy Use and Performance on Language Tests: A Structural Equation Modeling Approach*, Cambridge, Cambridge University Press.
- Richards, Jack C., 2005, *Language Teaching in Changing Times*, Plenaria, 12 de abril, Medellín, Cambridge University Press.
- Shohamy, Elana, 2001, *The Power of Tests- A Critical Perspective on the Uses of Language Tests*, Essex, Pearson Educational.
- Tejada Fernández, J., 1999, "La evaluación: su conceptualización", en: *Evaluación de programas, centros y profesores*, Madrid, Síntesis.
- Widdowson, Henry G., 1983, *Learning Purpose and Language Use*, Oxford, Oxford University Press.
- Yule, George, 1996, *Pragmatics*, Oxford, Oxford University Press.



LAS AUTORAS

** Magister en Inglés como Segunda Lengua. Docente Asistente de Tiempo Completo. Profesora investigadora de la Escuela de Idiomas, Universidad de Antioquia. Correo electrónico: clarias@idiomas.udea.edu.co

*** Especialista en Lenguas Extranjeras, Docente Auxiliar de Tiempo Completo. Fundación Universitaria Luis Amigó, Vicerrectoría Académica, Departamento de Idiomas. Correo electrónico: lilith39@epm.net.co

Miembros del Grupo de Investigación Acción y Evaluación en Lenguas Extranjeras (GIAE). Correo electrónico: giae@idiomas.udea.edu.co

