

AGUSTÍN YAGÜE BARREDO

Profesor titular de Español para Extranjeros en la Escuela Oficial de Idiomas de Barcelona-Drassanes, es asimismo es docente del Máster de Formación de Profesores de ELE de la Universidad de Barcelona (semipresencial) y del Máster del Instituto Cervantes – Universidad Internacional Menéndez Pelayo.

Ha sido Asesor Técnico del Ministerio de Educación y Ciencia de España en Nueva Zelanda y ha impartido numerosos cursos de formación de profesorado de ELE. Es igualmente autor de artículos y materiales didácticos para la enseñanza del español.

Ha participado en varios proyectos educativos de formación no presencial y de didáctica del ELE, entre ellos la creación de *redELE*, y diversos espacios web y listas de distribución para docentes, tanto de carácter internacional (Formespa) como de ámbito local (en Nueva Zelanda y Asia).

RESUMEN

El objetivo de este artículo es reflexionar acerca del concepto y el diseño de exámenes de español como lengua extranjera. Con criterios empíricos se subrayan las limitaciones técnicas de las pruebas y la pobreza de información que ofrecen a docentes y a candidatos sobre su actuación. El trabajo propone algunas alternativas que deben ser refrendadas en situaciones de aula.

PALABRAS CLAVE: exámenes, medición, Español Lengua Extranjera

TESTING TESTS. CONSIDERATIONS ABOUT THE MEASUREMENT TESTS FOR SFL

The present paper focuses on the concept and design of tests for Spanish as a Foreign Language. Their technical limitations -and the paucity of information they provide to teachers and candidates about students' performance- are highlighted through empirical criteria. The article proposes alternative models to be validated in the classroom.

KEYWORDS: testing, measurement, Spanish as a Foreign Language

Mi agradecimiento a mis colegas del Departamento de Español para Extranjeros de la EOI Barcelona Drassanes María José Hernández y Lourdes Miquel, por su amabilidad al "examinar" estas páginas antes de ser publicadas. Y a Glòria Feliu, con quien elaboré algunas de las propuestas aquí recogidas. Mi admiración por todas ellas por sus valiosas sugerencias.

CONSIDERACIONES PREVIAS: *LO QUE SÍ Y LO QUE NO*

El presente artículo¹ aborda de modo general algunas reflexiones sobre la concepción de exámenes de ELE en centros de enseñanza reglada, tanto los que pueda confeccionar un docente como los que son suministrados por las instituciones educativas. En esas reflexiones tiene una especial incidencia, además de la capacidad de medición, la capacidad de brindar información a docentes y aprendientes sobre la medición misma. Tal y como apunta Ortega (1995, 22):

mediante la realización de la correspondiente prueba, el profesor podrá disponer de información sobre el PROGRESO REAL de sus estudiantes, [...] o podrá ESTABLECER UN DIAGNÓSTICO SOBRE DETERMINADAS REALIDADES LINGÜÍSTICAS, sobre el nivel alcanzado en cierto momento respecto de alguna especificación de los objetivos, etc.; podrá, incluso, si está interesado en cómo se adquiere una lengua en el aula, sustentar o desechar hipótesis sobre los múltiples aspectos de tal proceso. Los estudiantes, por otro lado, deberían considerar la evaluación como lo que realmente es (o DEBIERA SER): instrumento que proporciona información sobre el proceso de aprendizaje, por ejemplo, y entre otras muchas cosas, MIDIENDO EL PROGRESO REALIZADO, SEÑALANDO PUNTOS DÉBILES, mostrando la adecuación entre lo que se hace en clase y los objetivos perseguidos, etc. (los destacados son míos).

Se formulan en primer lugar algunos trazos generales sobre el diseño y condiciones de administración de los exámenes para adentrarse después en las técnicas empleadas y sus limitaciones y, finalmente, proponer algunas alternativas que pueden en algún modo amortiguar parte de los problemas detectados.

El término evaluación es complejo: alude e incide –o puede aludir e incidir– sobre decisiones académicas y administrativas, de planes curriculares y configuración de sílabos, etc., a menudo abstractas y sin intervención directa del profesorado. En ese panorama complejo, los exámenes son sólo una herramienta –entre otras– para la evaluación, pero no conviene no olvidar que PARA MUCHOS ESTUDIANTES LOS EXÁMENES SON EL ÚNICO INSTRUMENTO DE EVALUACIÓN DE SU RENDIMIENTO / COMPETENCIA, ya sea en forma de pruebas de certificado (los DELE, pruebas de las EEOII...), exámenes nacionales o, en ciertos casos, los llamados "exámenes

¹ Este trabajo recoge y expande los contenidos presentados en el *XVIII Encuentro Práctico de Profesores de ELE* organizado por la Editorial Difusión (Barcelona, 18 y 19 de diciembre de 2009).

finales” ... Sólo en la elaboración de éstos últimos, circunscritos a un centro de enseñanza, pueden intervenir explícitamente los profesores implicados directamente en la docencia de los candidatos.

El objetivo de estas páginas es, precisamente, el concepto de diseño de exámenes, al margen de su trascendencia en un sistema de evaluación de orden superior, al margen –en parte- de los procedimientos de corrección / calificación, al margen igualmente de los contenidos objeto de “medición” establecidos por los desarrollos curriculares en que se imparte la docencia y por las escalas de referencia conocidas... Y, por descontado, al margen asimismo de esos “exámenes” –cualquiera que sea su perfil- articulados sobre la L1, la traducción y, casi exclusivamente, en baterías de formas gramaticales. Que haberlos, haylos. Como las meigas. Y dan miedo. También como las meigas.

Estas reflexiones, que pretenden ser abiertas, nacen de los sentimientos encontrados que aparecen cada vez que nuestros estudiantes deben afrontar una prueba de medición, o para entendernos mejor, un examen.

Es cierto que, más o menos, muchos de los exámenes validan intuiciones acerca de nuestros estudiantes o incluso datos más o menos fiables compilados a lo largo de un curso, y que, en definitiva, discriminan y clasifican con razonable certeza. Pero, ¿significa eso que nos podamos sentir satisfechos de su concepción y administración? A menudo, esquivamos esa pregunta...

Durante el desarrollo del examen, durante el proceso de corrección y tras el proceso de análisis de resultados nos asaltan las dudas... ¿por qué determinado estudiante ha fallado una pregunta de gramática que en las actuaciones de aula parecía producir correctamente? ¿dónde está el problema: en su actuación, en el diseño de la pregunta, en la falta de reconocimiento entre aquel exponente y este ejercicio?, ¿por qué un número significativo de estudiantes resuelve satisfactoriamente la parte estrictamente gramatical de una prueba pero sin embargo su producción oral o escrita no recoge en un modo aceptable esas presumibles competencias gramaticales?, algunas actuaciones de nuestros estudiantes –mecánicas, memorísticas, en algunos casos calificadas como de estratégicas-, ¿corresponden realmente a un aprendizaje significativo?, ¿por qué un estudiante se puede quedar prácticamente bloqueado en la interacción oral con un compañero/a en una situación de examen?, ¿qué hay de las diferencias culturales en el diseño de un examen, acaso un estudiante francés y otro chino pueden abordar unas mismas pruebas con idénticas garantías...?

Las preguntas son numerosas; las respuestas, posiblemente nuevas interrogantes, pero posiblemente se puede avanzar alguna conclusión, a modo de hipótesis: que LOS EXÁMENES -EL DISEÑO DE LOS MISMOS- ESTÁN LEJOS DE LOS ENFOQUES METODOLÓGICOS ACTUALES, lo cual genera una enorme distancia entre la realidad del aula, la docencia misma, y la medición de la competencia. El examen deviene una suerte de alienígena en unos planteamientos didácticos de aula presumiblemente sensatos. Por otra parte, LOS EXÁMENES NO PARECEN CUMPLIR

ADECUADAMENTE EN TODOS LOS CASOS (O NO CON GARANTÍAS SUFICIENTES) SU COMETIDO DE MEDICIÓN DE LA COMPETENCIA SINCRÓNICA DE UN APRENDIENTE.

Todos hemos tenido algún estudiante competente que ha fracasado en alguna prueba de este tipo. Lo hemos achacado a los nervios, a la tensión. Pocas veces a la concepción del examen, en sí mismo y en su relación con el candidato. En el otro extremo, hemos podido comprobar cómo algunos estudiantes se enfrentan con test clasificatorios gradados y “aciertan” preguntas que sobrepasan su nivel: el azar también cuenta, pero el diagnóstico puede ser erróneo. Es también cierto que algunos de los estudiantes que fracasan reflejan, ciertamente, un nivel de competencia insuficiente al esperado / establecido, pero el retrato que se ofrece de su competencia es romo, con pocos matices, en especial en las pruebas denominadas objetivas. Y en parte ello es debido a la ASIGNACIÓN DE PUNTUACIONES UNIFORMES: es casi tónica habitual que todos los ítems de una misma prueba tengan el mismo valor cuantificador, lo que deja al mismo nivel todos los contenidos, con independencia de la dificultad intrínseca de un ítem en concreto (se puede acceder a información estadística si el ítem ha sido pilotado con anterioridad), la complejidad general de adquisición de un contenido y las diferencias individuales (al menos las referidas a la L1 de los candidatos).

CONSIDERACIONES GENERALES. TÉCNICAS: ESTO ES LO QUE HAY

Las definiciones de evaluación son numerosas en la literatura de enseñanza de lenguas, pero no lo son tanto las referidas a exámenes². En su versión más “generosa” hablaríamos de *calificar y obtener información acerca del proceso de enseñanza-aprendizaje* (Bordón: 2000, citado por Pastor: 2009), lo cual, aun siendo deseable, no siempre se logra satisfactoriamente. En una versión más técnica, se podría adaptar el apunte de Buck (2001, recogido en Pastor: 2009), para aludir a la ejecución / resolución de tareas –por lo general simples- sin colaboración, lo que dejaría fuera las pruebas de interacción oral, algo que posiblemente sea aceptable ya que su diseño no comporta la intervención de la acción del estudiante en algún modo de escritura sobre papel.

No nos detendremos en esa descripción, pero no es banal subrayar que la resolución de tareas en solitario, sin cooperación, entra inmediatamente en colisión con el concepto mismo de lengua como actividad comunicativa humana...

Antes de adentrarnos en el concepto mismo de las pruebas, son procedentes algunos APUNTES SOBRE LA ADMINISTRACIÓN de las mismas.

² Las definiciones son en ocasiones difusas, sin límites precisos entre evaluación e instrumentos de evaluación, incluidos los de medición. Véase, sobre el particular, el capítulo 9 del MCER (2002): pág. 177 y ss. Los conceptos de fiabilidad, validez y viabilidad en él subrayados reaparecerán en este trabajo, en especial el primero de ellos.

En primer lugar, la DURACIÓN; en muchos casos los candidatos pueden invertir hasta cuatro horas en un único día para completar las diversas partes de estos exámenes... Pocos estudiantes habrán tenido la oportunidad de asistir a clases de esa duración: el agotamiento, la pérdida de concentración, convive con los exámenes; el esfuerzo cognitivo –repartido, disperso mejor, entre tareas diferentes- puede ser significativo; se debería determinar –no conozco demasiados estudios al respecto referidos al ELE- la incidencia de ese factor en el resultado final, con grupos muestra bien definidos sometidos a administraciones diferentes.

Se podrá argüir que muchos estudiantes están acostumbrados a exámenes largos en sus trayectorias académicas, pero eso no resta valor a la afirmación: no parece una cuestión de “costumbres”: tal vez una temporización diferente les hubiera dado mejores resultados, y en cualquier caso esos exámenes no se han realizado por lo general en una lengua distinta a la L1. Con todo, con frecuencia se parece entrever un perfil idealizado del estudiante de ELE (en manuales, en determinados diseños de cursos y ofertas formativas): el de una persona joven, culta, universitaria, inteligente, motivada... Una joyita, vamos. La realidad, sin embargo, es otra, o por lo menos más compleja y diversa que ésta. Y en cualquier caso, queremos entender que la duración de los exámenes es un problema importante para una medición pertinente. Y no vale decir que las condiciones son idénticas para todos: los estudiantes no son iguales. Otro asunto es que la concentración en el tiempo sea el procedimiento más cómodo o más económico para administradores y candidatos: una cosa no quita la otra.

En relación con lo anterior, convendría estudiar CÓMO ORDENAR LA SECUENCIA DE LAS PRUEBAS, Y EN RAZÓN DE QUÉ FUNDAMENTOS DIDÁCTICOS. La secuencia de los DELE dispone en primer lugar las pruebas de comprensión y producción escritas y a continuación la de comprensión oral y finalmente la de gramática y vocabulario. Los certificados de las Escuelas Oficiales de Idiomas de Cataluña proponen otra ordenación, que, por cierto, no coincide con la de las EEOOI de la Comunidad de Madrid... Si las decisiones son estrictamente internas, las razones pueden ser variadas: con frecuencia las tareas de comprensión auditiva no se acostumbran a poner en primer lugar (por aquello de si alguien llega un poquito tarde) y la parte de expresión escrita se suele emplazar al final de la prueba, como un apéndice.

El panorama parece algo errático, sobre todo si entendemos que comúnmente se quiere revestir a los exámenes con un barniz de prueba científica... Las decisiones deberían adoptarse con arreglo al perfil de los candidatos (si se conoce el grupo), de acuerdo con el esfuerzo cognitivo general requerido e incluso en función de los estilos de aprendizaje. Pero, sobre todo, de acuerdo con los rasgos concretos, específicos, de las pruebas que se van a administrar (que ése sí es un elemento conocido). Los redactores de exámenes no son máquinas perfectas y los productos finales no se muestran siempre enteramente compensados en sus partes. Conviene, por consiguiente, analizar esto antes de administrar los exámenes y ajustar el orden de acuerdo con las exigencias de esa prueba... Eso significa que la ordenación puede ser diferente en cada caso: si el equipo de redactores / examinadores estima que la prueba –esa prueba

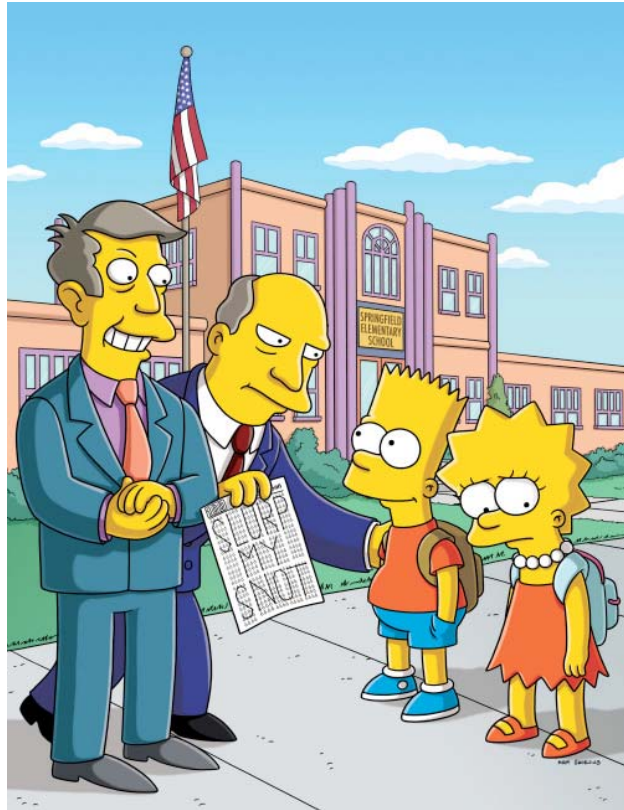
concreta- de comprensión escrita, pongamos por caso, exige un esfuerzo mayor que las restantes, tal vez sea oportuno ubicarla en primer lugar en esa ocasión, pero no en otras, o justo después de otra que resulte facilitadora en algún modo. En cualquier caso, lo que quiero subrayar es que el orden suele estar determinado de antemano, y suele estarlo con independencia de las pruebas mismas. No parece muy sensato. Es incluso razonable al menos valorar que el candidato disponga del conjunto de las pruebas, y que sea él mismo, en razón de su autonomía y responsabilidad de aprendizaje y estratégica, quien seleccione el orden (nuestros estudiantes no están demasiado acostumbrados a eso, de modo que sería pertinente prever preparación previa en este sentido). Lógicamente la prueba de comprensión auditiva es la más comprometida para ese tipo de secuenciación autónoma, si bien, como veremos más adelante, son posibles otras formulaciones –tal vez beneficiosas- que apuesten por un trabajo individualizado.

Otro aspecto, en absoluto banal pero que escapa a nuestras posibilidades de decisión, es EL CONFORT –LA FALTA DE CONFORT- DE LOS ESPACIOS EN QUE SE ADMINISTRAN LOS EXÁMENES. De muchas de las aulas de idiomas han desaparecido las mesas (presuntas barreras para la interacción oral y en cualquier caso limitadoras de variaciones del espacio de la clase) en favor de sillas de brazo con tablero. A excepción de la de expresión / interacción oral, el resto de pruebas exige que el alumno lea y escriba: la comodidad de esos asientos (la cuadratura del culo, que dijo alguien refiriéndose a asientos duros) es más que dudosa, pero no dudas en su calidad de espacio para escritura durante un tiempo prolongado, como decía. Las soluciones no están en nuestra mano, pero el problema y sus eventuales repercusiones sí inciden sobre los aprendientes.

No entro a tratar aquí otros aspectos significativos, como la iluminación o la calidad de la visualización / audición de los documentos audiovisuales de la prueba, pues las condiciones pueden variar sensiblemente. He apuntado simplemente el caso de las sillas –entiendo que bastante generalizado- como un punto de reflexión sobre la idea de confort, imprescindible para realizar la prueba en el entorno más satisfactorio.

Pero si las decisiones sobre mobiliario escolar suelen corresponder a las Administraciones educativas y por ello cualquier modificación es sumamente compleja, no lo debiera ser tanto LA CONFIGURACIÓN DEL SOPORTE –EL PAPEL, POR EL MOMENTO- DEL EXAMEN, en tanto que no comportaría una carga presupuestaria. Muchos de los exámenes, tanto oficiales como los creados por los departamentos de los centros docentes, suelen tener un diseño errático.

Dejando de lado por el momento otros aspectos, que reaparecerán más adelante (entre ellos, el muy relevante de la tipografía), una de las tendencias más acusadas es la de transferir las producciones / respuestas de los estudiantes hacia una hoja aparte, la llamada precisamente “hoja de respuestas”, generalmente presentada como una densa cuadrícula en la que los candidatos deben marcar con cruces la columna adecuada. Los errores son frecuentes. De hecho lo son también para cualquier nativo en cualquier tipo de test que se sirva de ellas, pero parece lógico suponer que se agraven en un



un trabajo en una L2. En otros casos, el espacio para las respuestas se sitúa incomprensiblemente en el reverso de la hoja de preguntas, de manera que el proceso de contestar exige que se gire repetidamente el papel de examen.

La disposición de dos o tres planos visuales (en algunos casos distantes, como sucede con la comprensión auditiva con soporte visual), todos ellos activos en el proceso, se ha revelado como perturbadora para los procesos de atención (Mayer, 1997, 2002). Las soluciones no son demasiado complejas, incluso con conocimientos muy rudimentarios de procesadores de texto: el objetivo, aunar enunciados de preguntas y respuestas en un mismo plano gráfico, y como mucho dos; evitar en cualquier caso que, por ejemplo, un trabajo de comprensión escrita tenga el propio texto de lectura, la hoja de preguntas y, finalmente, la de respuestas. Recuérdese, en nuestra silla de brazo abatible no caben, expuestas para una operación simultánea, esas tres hojas. Se caen al suelo, sin remedio.

EJEMPLO. A partir de la lectura, se puede afirmar que el autor del texto

- a) No conoció personalmente a Kevin Carter
- b) Conoció a Carter en Nueva York
- c) Estuvo en contacto con Carter en Sudáfrica

☒ C

1. Según el texto y en relación con la fotografía de la niña y el buitre, Kevin Carter

- a) no se siente satisfecho con las fotos que ha hecho
- b) piensa que ganará algún premio importante con las fotos que ha hecho
- c) espera 20 minutos para hacer una foto

☐

Ejemplo de hoja de respuestas en el mismo plano visual de las preguntas

El precio, la renuncia a parrillas de corrección superpuestas o correcciones automatizadas, que desde luego simplifican el trabajo del docente en lo que es una tarea extremadamente aburrida. Lo cual no está nada mal.

Sin abandonar estas consideraciones generales, nos adentramos algo más en el diseño mismo de los exámenes, nuestro objetivo, entendiendo, con Ortega (1995, 23) que

una prueba configurada correctamente, consistente y válida respecto de los objetivos de un curso de lengua y de la práctica docente seguida durante el mismo, es -sin pretensión de caer en radicalismo alguno- un instrumento de innegable utilidad con el que valorar y tomar decisiones sobre docencia y aprendizaje. [...] La finalidad de todo proceso de aprendizaje / docencia es -y nadie que esté en su sano juicio podrá negarla- aprender / enseñar algo, y no hacer pruebas; es un dislate confundir los objetivos docentes o de aprendizaje de un curso con las pruebas que puedan o deban tener lugar durante o al final del mismo.

Y así establecer qué es, o puede ser, “una prueba configurada correctamente, consistente y válida”.

Se suele distinguir entre PRUEBAS SUBJETIVAS Y PRUEBAS OBJETIVAS (se entiende que los adjetivos califican a la “corrección” de las pruebas). Las primeras responden a una producción libre o semicontrolada del candidato, a partir de una(s) instrucción(es) cuya arquitectura general acostumbra a ser conocida por el mismo. Las segundas, se articulan como baterías de preguntas que exigen una selección entre dos o más alternativas de posibles respuestas, unas sin supuestos (opción entre verdadero o falso), y otras con eventuales respuestas propuestas (selección múltiple, en sus diversas –y limitadas- variedades).

Como es natural, la atribución de “subjetivo” u “objetivo” no puede recaer exclusivamente en la corrección. El principal ingrediente de la “subjetividad” descansa en la elección y la selección de los textos (escritos o auditivos) y en las preguntas (y a veces en las respuestas) asociadas a ellos. Ésa es la pregunta del millón de dólares, ¿qué preguntar?, ¿con qué textos o documentos? ¿por qué? ¿es posible encontrar los textos o documentos más adecuados, o debemos “conformarnos” con muestras aproximadas, razonables?). La casuística es muy variada (hay exámenes, textos, ítems “buenos” y los hay “malos”, entendidos ambos adjetivos en su valor más rudimentario, de sensatez inmediata). En algún modo, las denominadas pruebas “objetivas” tienen mucho de “subjetivas”...

No nos es posible aquí analizar demasiadas muestras reales³ (en algunos casos no son fáciles de conseguir y mucho menos de divulgar, a pesar de haber sido

³ La selección de muestras institucionales, oficiales, que recoge este trabajo responde a diversos criterios. Uno de ellos es su actualidad (se ha preferido en todo momento pruebas recientes: la mayoría de ellas corresponde a 2009). Otro, su capacidad para ilustrar algunas de las carencias que se apuntan en este artículo, lo cual no significa en absoluto que no existan ítems o conjuntos

empleadas con estudiantes: muchas instituciones las preservan, tal vez para ser reutilizadas): no solo nos llevaría mucho tiempo sino que además los resultados obtenidos no tendrían por qué ser necesariamente extrapolables a otras muestras. Nos referiremos aquí, por consiguiente, a fenómenos constantes del diseño de exámenes...

Una de esas constantes es sin duda el REPERTORIO LIMITADÍSIMO (Y A MI JUICIO LIMITADOR) DE LAS HERRAMIENTAS DE DISEÑO DE PRUEBAS "OBJETIVAS". En efecto, no van mucho más allá de los ya apuntados arriba:

a) seleccionar si un enunciado escrito, referido a su vez a un documento, escrito o audiovisual, es verdadero o falso.

b) seleccionar entre, por lo general, tres o, como máximo, cuatro opciones, también escritas, referidas a su vez a un documento, escrito, auditivo o audiovisual.

c) variación de la anterior, en especial en las pruebas de competencia gramatical y o léxica, selección en la que no se ofrece un enunciado como tal, sino únicamente diversas formas –gramaticales y/o léxicas-, de las que tan solo una (en el mejor de los casos, cuando el examen no presenta ambigüedades) es la pertinente en el contexto propuesto (frase o texto, aunque tal distinción no es significativa en muchos casos, como veremos).

d) como variación también de la anterior, sobre todo en las pruebas de comprensión escrita, la selección de fragmentos para asociar (completar, titular) párrafos o textos breves... A modo de perversa oferta del súper: si falla usted una, fallará dos: garantizado⁴.

Otras opciones (como la ordenación de los constituyentes sintagmáticos de un enunciado, la detección de errores, determinadas transformaciones gramaticales...) han dejado, por lo general y afortunadamente, de emplearse en los exámenes de ELE, por lo que no aludo a ellas.

Como es lógico, pueden existir otras alternativas para acercarse a un texto escrito u oral en situación de examen (algunas de ellas las abordaremos más adelante), pero lo cierto es que no menudean en las pruebas publicadas o conocidas. Las cuatro agrupaciones anteriores aparecen refrendadas como "fiables" por los estudios de psicometría y se mantienen sin apenas desgaste desde hace años. Sin embargo, si aplicamos algo de sentido común, veremos que esa fiabilidad puede resultar pobre y además escasamente informativa para

construidos razonablemente. Para los objetivos de este trabajo, las primeras, la que presentan deficiencias, logran, a mi juicio, suscitar una reflexión más rica que los segundos; eso explica su inclusión, no la voluntad de ofrecer una visión sesgada del panorama.

⁴ La excepción, proporcionar un excedente de opciones. Queda por calibrar el rendimiento y eficacia de esta formulación, más allá de la evidente minimización del error por exclusión.

docentes y discentes, y que nuestras formas de aproximarnos a la didáctica de lenguas segundas han evolucionado sensiblemente.

Otra de las constantes derivadas de lo anterior es el PREDOMINIO DEL TEXTO ESCRITO COMO MEDIO EXAMINADOR EN CUALQUIERA DE LAS PRUEBAS (a excepción, por definición, de la de interacción oral, aunque en ella también hay –al menos- instrucciones escritas). Además de marginador para estudiantes con lenguas no próximas al español, el procedimiento parece ya en primera instancia antinatural: no hay razón ni técnica ni pedagógica que impida que alguna(s) de las preguntas de, por ejemplo, comprensión oral se formule(n) también oralmente, aunque posiblemente eso suponga modificar los posibles diseños de articular preguntas y respuestas.

Por otra parte, en los casos que corresponde, EL ENUNCIADO ESCRITO AL QUE SE VINCULAN LAS PREGUNTAS SUPONE UNA REFORMULACIÓN DEL DOCUMENTO –ORAL O ESCRITO– ORIGINAL, teóricamente el que el estudiante debe abordar y entender. Tal planteamiento supone la coexistencia de dos discursos simultáneos que generan una sobrecarga cognitiva singular. Si no se entienden los enunciados de las preguntas, “no se entiende” el texto (a pesar de que el texto sí puede haberse entendido). Por lo demás, el redactor de pruebas suele no abordar literalmente los contenidos del texto primero, sino redactar su enunciado con paráfrasis y, sobre todo, sinónimos... ¿Quién no ha tenido estudiantes que se han quejado de que las preguntas –no el texto ofrecido– eran “muy difíciles”? De hecho, casi todos los profesores intentamos aconsejar a nuestros estudiantes que lean atentamente las preguntas... Raro, raro: el consejo lógico sería que leyeran o escucharan con mucha atención los textos objeto de comprensión y que desplegaran las estrategias apropiadas... La competencia estratégica de afrontar exámenes de lenguas, ¿es realmente una competencia que forme parte de los intercambios comunicativos entre los seres humanos?, ¿hasta qué punto debe serlo y en qué modo se valora (recuérdese que las instrucciones son una parte no menor de los exámenes mismos y de la preparación de éstos)?

En efecto, buena parte del trabajo de comprensión se reduce a un trabajo de vocabulario (entendido en un sentido amplio), que requiere del estudiante una significativa competencia léxica, a la que hay que añadir los matices perdidos o alterados en la nueva formulación con respecto al original. Esta concepción hace que la carga de la comprensión resida en unidades inferiores –el sintagma o la frase–, más que en el propio texto como conjunto. El proceso es especialmente agotador y angustioso cuando se funda en la recuperación auditiva, efímera por naturaleza, con poco margen para desarrollos estratégicos que sí tienen mayor cabida en la comprensión escrita, con espacio para inferencias, analogías y otros mecanismos cognitivos.

a) VERDADERO / FALSO, ¿ES ÉSA LA CUESTIÓN?

Retomando el asunto de la fiabilidad, no parece que éste sea sostenible en los enunciados (con el doble discurso apuntado) fundados en respuesta tipo verdadero / falso: el 50% está garantizado de antemano.

En los casos en que estos sistemas de doble opción son utilizados, por ejemplo, en los test empleados en el mercado laboral, los cómputos finales –son por lo general cuestionarios extensos- suelen incluir un índice corrector con el que paliar el grado de aleatoriedad... No hacemos eso en nuestros exámenes de, por ejemplo, comprensión auditiva; no eliminamos, pongamos por caso, dos de las diez preguntas (¡que tanto nos ha costado idear!) para objetivar la prueba (que por lo demás es una de las de corrección objetiva)... O es verdadero o es falso, o está bien o está mal. En ninguno de los casos la respuesta nos resulta informativa por sí misma: el candidato, ¿ha entendido o ha acertado? A veces ni siquiera lo es la pregunta, en especial cuando está se formula con alguna forma negativa.

En algunas de las pruebas realizadas con mis estudiantes, he introducido algún elemento delimitador para validar este tipo de respuestas. El planteamiento es muy simple y puede ser más o menos automatizado en los casos de comprensión escrita; TRAS MARCAR LA OPCIÓN “VERDADERO” O “FALSO”, EL ESTUDIANTE DEBE ANOTAR LA LÍNEA –O LÍNEAS- EN QUE APARECE LA INFORMACIÓN SOLICITADA (los textos siempre disponen de numeración en el margen izquierdo). Si la pregunta se refiere a un documento audiovisual, la gestión es más compleja, y ha de ser particularizada en cada caso, y puede incluir una pregunta doble, abierta o cerrada: la justificación de la elección de verdadero o falso queda arropada con otra respuesta de redactado breve o de selección múltiple (lo dice la mujer / lo dice el hombre; lo dice al principio / lo dice antes de la comida, o similares). Como digo, la gestión es más compleja, en especial el redactado de los enunciados para que no resulte muy extenso (puede ser adecuada en fragmentos breves, con pocos ítems). En cualquier caso, no tengo constancia de que estos planteamientos hayan sido especialmente perturbadores para los estudiantes: en muchos casos la respuesta aparece condicionada por la memoria auditiva, y esa pregunta “doble” puede contribuir a la recuperación de la misma.

Otra opción que he podido experimentar para este tipo de ítems de pareja excluyente es la de INCLUIR ALGUNA FOTOGRAFÍA EN EL TEXTO O EN LA HOJA DE PREGUNTAS QUE ACTÚE COMO ÍTEM, esto es, rebasando el concepto de simple ilustración (también valioso, incluido su componente afectivo, y, por ejemplo, plantear si la fotografía guarda (*sí* o *no*, en vez de *verdadero* o *falso*) relación (directa) con el texto, un procedimiento que suele resultar amable y que evita al menos uno de los enunciados reformuladores del texto original.

DÍME CÓMO BESAS...

Salud a pedir de boca

Un beso en la boca es, científicamente, una especie de elixir. Sí, está comprobado. Existe, incluso, una ciencia, la filematología, dedicada a estudiar los besos. Además, afirman los expertos, las sustancias químicas que la saliva contiene actúan como una especie de detector de compatibilidad entre las posibles parejas.

5

Cuando besamos, la producción de cortisol, conocida como "la hormona del estrés", disminuye y la oxitocina - también llamada "molécula de la monogamia", "molécula de la confianza" o "sustancia del amor"- aumenta. Esencial en el desarrollo del afecto y de la necesidad de cuidarse mutuamente, la oxitocina circula a todo tren por nuestra corriente sanguínea en el momento del beso. De hecho, siempre ha sido conocida por su poder en la maternidad, ya que fortalece los vínculos entre madres e hijos al ser una hormona que se libera en grandes cantidades durante el parto así como en la lactancia de los bebés.

10

15

20

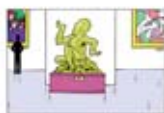


La fotografía que ilustra el texto (no incluida en la versión publicada del mismo) sirvió como pregunta de *sí o no* en un trabajo de comprensión lectora para estudiantes de B1+. (*¿Está la fotografía relacionada con el texto?*). A pesar del título general del texto, el candidato debía hallar diversas informaciones repartidas en el fragmento (besos, bebés), valorarlas, para, finalmente, determinar la opción adecuada. En este caso se les solicitaba también que anotaran la(s) línea(s) que le permitía justificar su elección

Este uso de las fotografías está por lo general poco explorado en los exámenes, pero su rendimiento suele ser excelente con una adecuada concepción. Los exámenes A1 del Instituto Cervantes incorporan algunas imágenes (no siempre de la calidad y la precisión deseables), sobre todo vinculadas al trabajo léxico.



A



B



C



D



E



F



G



H



I

Propuesta de trabajo de comprensión auditiva (emparejamientos) de la convocatoria de mayo de 2009 del DELE A1. La calidad de las imágenes es irregular, y también lo es su asociación: la imagen C no resulta descriptiva o la imagen I (en la que aparece el dibujo de un hombre) se asocia con una voz femenina en el audio.

5. Susana va a llevar a la cena...



a)



b)



c)



d)

Propuesta de comprensión lectora aparecida en la misma convocatoria citada. El texto inicial (un mensaje de correo electrónico con una invitación a una fiesta) alude explícitamente a un *vestido verde*. Al tratarse precisamente de un input escrito, las opciones de respuesta eluden la escritura para evitar el reconocimiento. La opción es ágil –posiblemente más adecuada para comprensión auditiva- e impide la aparición de sinónimos y parafraseos –comprometidos con el nivel de los candidatos-, pero no parece ir más allá de un *matching cards* de vocabulario, insuficiente para un trabajo de comprensión lectora.

También, como veremos, en otros tipos de pruebas. En el caso de la comprensión oral, como estimulador de la memoria auditiva (previsiblemente con respuestas más inmediatas, del tipo ¿quién ha dicho tal enunciado –literal o no-? a lo que el alumno debe seleccionar una fotografía (capturas de pantalla) entre tres o propuestas similares, y desde luego –nuevamente con capturas de pantalla- como focalizadoras de expresiones faciales y actitudes. Incluso en la gramática; si en clase intentamos en ocasiones sistematizar diversos usos de la lengua con representaciones gráficas (por ejemplo, las preposiciones, diversos matices verbales...), no hay razón para no recuperar algunas de esas representaciones en un examen. Lo dicho: el examen no debe ser un alienígena en el proceso de enseñanza / aprendizaje.

Con estos simples añadidos, es posible conseguir que este tipo de respuestas queden, por un lado, legítimamente validadas y que, por otro, sean poderosos informantes (en especial en los casos de error) para conocer, si es el caso, qué constituyentes en el texto pueden haber inducido al error).

De nuevo en nuestra búsqueda de constantes, el planteamiento anunciado se distancia de una ESTRUCTURA UNIFORME DE LA FORMULACIÓN DE ÍTEMS: TODAS LAS SERIES VINCULADAS A UN DOCUMENTO SON DE LA MISMA CLASE: series de verdadero / falso o bien de selección, en alguna de sus variedades. En el caso anterior, el híbrido resultante (de bajo riesgo, incluso para la corrección) reporta una interesante validación de series en las que el azar pesa en exceso, de modo que tal inclusión se muestra eficaz. No comparto la idea de que un modelo uniforme, único, de ítems constituye necesariamente una “protección” para el candidato.

Por otra parte, todos sabemos por experiencia lo complejo que resulta hallar un documento que cumpla las expectativas de aprendizaje / medición depositadas en él; una vez hallado, la explotación del mismo, en este caso como examen, no lo es menos (en especial en los documentos auditivos, de casi imposible modificación / adaptación). Inmersos en la explotación, es sin duda un problema encorsetarla en una única serie de ítems: algunos de los contenidos encajan de forma natural como selección múltiple (por ejemplo, en un documento oral, un trabajo de discriminación fonética); otros, en cambio, se desarrollan mejor –son más eficaces– con otros planteamientos.

Como apuntaba arriba, puedo afirmar que no he detectado problemas severos en los trabajos realizados con mis estudiantes, a lo sumo algún ajuste de tiempo en los ítems que incluyen producción abierta o semiabierta, que se deberá prever y prevenir.

b) SELECCIÓN MÚLTIPLE, EL ABC (ELIJA UNA) DE LOS EXÁMENES

El concepto de fiabilidad es aquí más sólido, sobre todo si las opciones contemplan cuatro alternativas, aunque la inclusión de esa cuarta opción es de elaboración costosa en numerosos casos. Tal vez ésa es la razón por la que la mayoría de exámenes se organizan con tres supuestos⁵, a modo de técnica “trilera”. De ellos, al menos uno suele actuar como “distractor⁶”, de manera que, aun sin poder generalizar, su capacidad de discriminación es “diferente”. Y así, no siempre es real y efectiva la propuesta de tres o cuatro opciones.

Estos repertorios pueden adoptar, por lo general, dos formas. La primera, con enunciados completos; la segunda, predominante en el trabajo gramatical y de vocabulario, sin enunciados, incluye únicamente las formas propuestas para completar el enunciado original ofrecido (no hay en este caso reescritura del texto original): nuestro apartado c).

Como es natural, la sensatez didáctica de este tipo de pruebas se funda en buena medida en la calidad de los redactados y de las opciones propuestas. No son infrecuentes los errores por parte de los redactores de pruebas, a quienes se les desliza más de una opción correcta (o al menos demasiado ambigua)... , que los estudiantes indefectiblemente suelen detectar. Verbigracia:

⁵ Uno de los aspectos más olvidados de éste y otros modelos es precisamente el concepto mismo de selección. La medición se concentra en la corrección del ítem seleccionado y desatiende la capacidad de información que puede brindar la selección de ítems incorrectos, si el repertorio propuesto tiene capacidad para brindar información.

⁶ El papel de los distractores está poco definido. Seguramente, como me apuntaba Lourdes Miquel, el número y la función de los distractores deberían estar determinados por el asunto que aborde el ítem: no pueden recibir el mismo tratamiento los pronombres personales que el contraste *ser/ estar*.

56. – Pero, ¿qué te contestó Javier?
– Que _____ que terminar el balance antes del próximo martes.
a) había tenido
b) tenía
c) tuvo
d) ha tenido

Pregunta de Gramática y vocabulario correspondiente a la convocatoria de los DELE (B2) de mayo de 2009. Ejemplo apuntado por Pablo Martínez Gila y Lourdes Miquel.

La redacción de propuestas absolutamente unívocas es seguramente posible (y ardua), pero en cierto modo trasluce en determinados casos unos modelos de lengua demasiado rígidos... Si se trata de enunciados "redactados" y si las opciones son consistentes entre sí –no extremadamente dispares–, el planteamiento puede restar interesantes capacidades de intervención de los candidatos, por ejemplo, y especialmente en las pruebas de comprensión escrita, la de interpretación (del texto y de la pregunta): algunos estudiantes intentan aplicar una lógica de comprensión fundada en su comprensión del texto y en su conocimiento del mundo: algunas de sus formulaciones que se sirven de inferencias e implicaturas son enteramente sensatas; sin embargo, frente a esas respuestas, la acción del examinador es ceñirse a la extrema literalidad ("eso no lo dice el texto"). En situación de examen, se cierra, y con portazo en las narices, una de las acciones que los hablantes hacemos de forma natural en la recepción de documentos escritos o sonoros, y que además solemos propugnar en el aula... Recuérdese por lo demás que algunas de las interpretaciones / explicaciones que los candidatos esgrimen para la selección de una respuesta concreta reflejan un grado de comprensión excelente o aceptable del texto y del enunciado; sin embargo, persiste el concepto penalizador de la literalidad.

Como se ha apuntado arriba para las opciones de verdadero / falso, también aquí es adecuado, en los casos en que sea posible, incluir explícitamente el añadido de localización de la información en el texto original indicando el número de la línea en que se encuentra. Las ventajas son al menos tres: se valida la fiabilidad de la respuesta, se promueve la reflexión del candidato y se fuerza al redactor de la prueba a constatar nuevamente la pertinencia de la opción seleccionada (más allá de su intuición de hablante nativo o competente, si la información que debe localizar el candidato es adecuada a la competencia prevista).

Este tipo de selección múltiple que comporta la redacción de un nuevo enunciado exige, como decía, un proceso de descodificación doble: el documento original y el enunciado mismo, con cierta tendencia a redactados con mayor carga léxica, que eluden la referencia literal al original generando opacidad al mismo (que será sin embargo el exigido como validador de la respuesta correcta). Con frecuencia el redactado de estos enunciados suele ser

breve, como si la pregunta no quisiera acaparar protagonismo en el ejercicio, pero el resultado puede ser confuso: los candidatos leen y releen las preguntas, convencidos –como aseguran– de que tienen “trampa”. Como procedimiento para el análisis del proceso de enseñanza / aprendizaje no es como para tirar cohetes que digamos...

El cuidado en el redactado debe ser extremo en estos casos (y es extremadamente difícil), y no debe eludir, como acostumbra a hacer, la inclusión de complementos (singularmente causales) e informaciones accesorias que permitan validar una de las opciones en detrimento de las restantes... La técnica, en el fondo, debería responder a una de las estrategias de mediación que recoge el *MCER*, la del parafraseo, que propicia la comprensión, esto es, no se orienta a entorpecerla. Muchos redactores de pruebas y no pocos docentes entienden esas informaciones accesorias como “pistas” que facilitan, por consiguiente, la elección de la respuesta correcta. La actitud penalizadora subyace en alguna de estas concepciones: intentar determinar lo que el candidato no sabe, mucho más de lo que sí sabe. Y el objetivo no es la penalización, sino la medición, la constatación, la detección, la observación, el análisis..., conceptos constructivos todos ellos.

C) ¿ES ÉSTA LA BUENA?

El sistema de selección múltiple sin redactado de enunciados es uno de los más utilizados y de los considerados como más fiables (véase, sobre el particular la nota 5, arriba). Sin entrar a cuestionar esta valoración, lo cierto es que su capacidad de medición presenta algunos problemas. El primero, el carácter antinatural que tiene la lectura de un texto trufado de huecos; este tipo de ejercicios, con buen rendimiento en actividades capacitadoras, dista mucho de ser una tarea completa. En segundo lugar, suele ser el terreno mejor abonado para los deslices de los elaboradores de pruebas según velamos en un ejemplo anterior, en especial cuando el objeto de atención es el sistema verbal. Si la correspondencia entre las opciones es muy lejana, la ambigüedad disminuye, pero también la capacidad de discriminación. Por otra parte, este tipo de trabajo se muestra limitado como informante, a menos que el repertorio de opciones sea próximo y el cotexto, muy sutil... Y este caso, para incrementar su valor informativo respecto a la interlengua del aprendiente, tal vez sería necesario PROCURAR LA REAPARICIÓN, EN OTRO MOMENTO DE LA(S) PRUEBA(S), DEL FENÓMENO QUE SE DESEA CONSTATAR. En efecto, en el proceso pueden intervenir el de selección y/o el de descarte; el primero, como decía, debe ser ratificado; el segundo es bajamente informativo: simplemente no sucede en las actividades comunicativas humanas habituales.

En otro orden de cosas, y siguiendo con este tipo de pruebas, la prevención frente ambigüedades y la presencia de más de una respuesta válida corren el riesgo de asentarse sobre MODELOS DE LENGUA ENCORSETADOS, raquíticos e insuficientes. Subyace en todo ello, un error de bulto: la incapacidad de DIFERENCIAR CONOCIMIENTO GRAMATICAL Y COMPETENCIA GRAMATICAL. Y en efecto, conviven aquí ciertas contradicciones: que la medición de lo gramatical

comúnmente se asienta en textos escritos (si bien en cierto modo dissociada de la comprensión lectora) y no en otros actos de lengua, y solo limitadamente en muestras productivas (véase, más abajo, el apartado de dedicado a la expresión escrita); que la atención al significado es predominantemente baja en este tipo de trabajos; y que la unidad de atención se sitúa mayoritariamente entre el sintagma y la frase (las raras excepciones son los trabajos con conectores, por lo general con un tratamiento muy mecánico, de orden estrictamente léxico (Roser Martínez, 1997). Todos ellos presupuestos lejanos a los planteamientos del aula...

No resulta fácil modificar el panorama descrito, entre otras cosas por lo arraigado de su uso. De nuevo las propuestas de opciones –y, de ser posible, su reaparición- resulta capital. En este sentido, tal vez sería interesante ensayar –no lo hecho nunca en situación de examen, pero la gestión no resultaría demasiado compleja con las prevenciones previas necesarias con anterioridad a la prueba- la posibilidad de seleccionar una opción (preferiblemente entre cuatro) y que el estudiante pudiera determinar si ha dudado o no de su respuesta. En el caso de que haya dudado, podría SELECCIONAR UNA SEGUNDA RESPUESTA ALTERNATIVA (la primera quedaría marcada con un 1 y la segunda con un 2, o algún procedimiento parecido). Este planteamiento, de uso esporádico y que podría tener algún tipo de cuantificación, sería capaz de proporcionar una medición más adecuada: no es el mismo tipo de error seleccionar, por ejemplo, un tiempo de pasado inadecuado (pretérito imperfecto / pretérito indefinido) que, por ejemplo, una forma verbal en futuro en ese mismo contexto..., por poner algún caso algo extremo. La “segunda opción” sería, además un valioso instrumento de información para el análisis de errores. Como es natural, el cuidado en la selección de opciones debería estar fundada en ítems gramaticales o léxicos próximos, incluido el papel de la opción “distractora”; de otro modo la información obtenida sería irrelevante.

Ello debiera permitir igualmente afinar el redactado, pero sobre todo la redacción de ítems, en consonancia con los contenidos objeto de medición. Concepciones más certeras vinculadas a lo que las investigaciones sobre adquisición nos aportan. Así, por ejemplo, si aceptamos la validez de *la estrategia de procesamiento léxico*, según la cual

los aprendices parecen prestar más atención a los elementos léxicos -los adverbios- que a la morfología verbal cuando procesan o interpretan el tiempo -presente, pasado, futuro- de un enunciado, debido a que los adverbios son mucho más perceptibles en el input que los morfemas verbales y, por lo tanto, contienen un valor comunicativo superior para el aprendiz (VanPatten 1984, 1985, tomado de Cadierno, 2010).

deberíamos concluir que algunas mediciones propuestas pueden quedar alteradas si el objeto de medición es precisamente la morfología verbal, lo que no significa que los mismos ítems sean adecuados con propósitos semejantes o complementarios (tanto gramaticales como –quizá no previstos- léxicos).

En cualquier caso, QUEDA PARA EL DEBATE SI LA MEDICIÓN DE CONOCIMIENTOS GRAMATICALES Y LÉXICOS DEBE ABORDARSE DE MODO AISLADO (no existe, por ejemplo, “prueba gramatical” en los exámenes de certificado de las EEOOI de la Comunidad de Madrid), en ejercicios específicos, o si debería ser la suma de diversos ítems contenidos en otras pruebas, especialmente las productivas. Esto es, puede sonar razonable disponer de una medición –llámese calificación gramatical y léxica, pero, ¿acaso no es posible repartir una serie de preguntas específicas entre las restantes pruebas hasta conseguir un número suficiente de ítems (quizá algunos con puntuación doble) para ofrecer esa medición...? Me inclino a creer que sí, y con considerables ventajas pedagógicas, ya por el simple hecho de que esas muestras gramaticales y léxicas se hallan en uso y pueden activarse también en procesos productivos (rastreo de determinados exponentes, o incluso reflexiones, en la interacción oral y la expresión escrita; más tarde regresaremos sobre este punto).

La complejidad del asunto es sobre todo técnica, en el momento de la calificación, pues comporta repescar anotaciones diversas en algún caso gradadas o dobles para configurar una medición / calificación propia... Un lío, sin duda. Pero lo que resulta insostenible, por bochornoso, es despachar la medición de la competencia gramatical y/o léxica con una docena y media o dos de ítems, resuelta como cruces en una parrilla y cuantificados con una plantilla. No es de recibo. En la actualidad, la inversión gramatical en el aula es considerable y sensata en su presentación y en su ejercitación... Su medición también ha de serlo.

Y lo ha de ser con otros criterios. Se percibe de alguna manera por parte de los docentes que el examen de gramática es el “más objetivo” de todos, el que proporciona datos fiables sobre los conocimientos del aprendiente de una L2, el inapelable... También los estudiantes se suelen mostrar siempre inquietos ante esta prueba... ¿Por qué temer la aparición del subjuntivo en este examen y no expresar preocupación excesiva por la eventual aparición en las lecturas, en las audiciones o en la interacción con un compañero?

d) CADA OVEJA CON SU PAREJA

El uso de este procedimiento es francamente limitado en la redacción de exámenes, circunscritos habitualmente a la comprensión escrita (en el caso de los DELE, en el nivel avanzado). Es una variación de la selección múltiple, en la que el candidato debe asociar fragmentos (en ocasiones títulos) con determinados textos breves o párrafos, discurso con entidad propia en definitiva. El planteamiento suele ser por lo general sensato: las opciones que se ofrecen acostumbran a formar parte del texto original: son, pues, enunciados no redactados por el elaborador de la prueba, y la resolución exige la comprensión general del texto (siempre breve, como decía), aunque en algunos casos se pueda requerir comprensión fina en determinados pares: la acción requerida del estudiante rebasa en ocasiones la unidad frase. En el polo opuesto, la mencionada “oferta envenenada”: un error en una asociación comporta automáticamente al menos otro. Y si bien esta formulación puede dar lugar a

autocorrección (fenómeno que precisaría una investigación propia), en ocasiones las cosas se complican si la intervención rebasa el par de opciones (las variaciones con tres o más elementos pueden perturbar el trabajo y generar un consumo importante de tiempo). Otros problemas asociables pueden ser, en ciertos casos, una carga excesiva de lo léxico (cuando se trata de asignar títulos), menos significativa en los casos de incorporación de fragmentos al documento primero, donde pueden intervenir elementos morfosintácticos, semánticos y textuales, en un sentido amplio. Finalmente, posiblemente el principal inconveniente de la prueba es la localización de un grupo de textos lo suficientemente próximos en lo argumental y en ciertos aspectos lingüísticos como para que las asociaciones resulten discriminadoras. La unidad la puede proporcionar la entrevista, al menos ése es el género que menudea en las pruebas, aunque en ocasiones los resultados son insólitos.

Columna A	Columna B
1 Dicen que no hay mejor fotografía de uno mismo que la autofotografía. Pero, ¿no piensas que seguramente se engaña uno a sí mismo más que al fotógrafo?	A Creo que empecé haciendo poesía visual y poco a poco me di cuenta que me gustaba contar historias y que, en cualquier lugar, en cualquier momento (hasta entre las sábanas), con un lápiz y un papel podía crear mundos, ilusiones, imágenes, emociones...
2 ¿Por qué has decidido llamarte Ouka Leele teniendo un nombre tan bonito, Bárbara?	B Es el cartel de una película del director Rafael Gordón, que es el que está haciendo una película sobre mí. Está inspirado en la película y mi sorpresa es que, cuando lo vi publicado, tanto en blanco y negro como en color, tiene una fuerza increíble y funciona muchísimo como cartel.
3 Además del libro de poesía, tienes pendiente el estreno de una película documental. ¿Podrías darnos más información sobre ella?	C Me interesan las imágenes aunque midan dos centímetros. Eso, como tú dices, es solamente comercial. Yo hago una diferencia entre el arte ornamental que podría estar en tiendas de decoración y el arte que hace temblar el corazón, el arte que cura y que nos hace sentir que hay algo más que la realidad pura y dura.
4 También has empezado a escribir poesía. ¿Cuánto de tus fotografías hay en tus versos? ¿Qué es lo que la poesía puede expresar que no te cabe en una foto (y al revés...)?	D Pues no sé, no soy consciente. Lo que sí sé es que algunas personas me han dado las gracias porque al ver mi obra se han sentido libres para hacer la suya. ¿Quizá la libertad?
5 “..Herida como la niebla por el sol.” Cuéntanos cuándo el cielo te sonríe con dientes de luna.	E Porque al principio me quería esconder. Quería crear un misterio donde no se supiera ni la edad ni el sexo ni la procedencia ni la época y, al crear ese misterio y esa aureola alrededor de mi obra, hacerla más deseable.

Muestra del DELE superior (excepcionalmente corresponde a 2007) en el que se observan asociaciones insólitas, incluso para el nivel: obsérvense los textos 1 y, singularmente, el 5, cuyas asociaciones se hallan en el reverso. Por lo demás, para completar correctamente el ejercicio basta con recuperar una única pieza léxica repetida en la columna de respuestas de la entrevista (subrayada además con algunas respuestas de “sí” o “no”). El procedimiento de repetición / recuperación es absolutamente normal en el intercambio de preguntas y respuestas que articula una entrevista, pero comporta aquí que no sea necesario leer / entender el texto para resolver la actividad de comprensión escrita.

No me consta que se hayan explorado demasiado en situación de examen otras posibilidades vinculadas, por ejemplo, a la comprensión oral (que sí tienen cabida en los manuales): vincular determinadas informaciones escritas a los distintos hablantes de una conversación grabada, por ejemplo. La administración de este tipo de prueba no es complicada, y su única limitación a priori es la capacidad de retención auditiva, que puede ser paliada con enunciados redundantes en cierta medida o, incluso con apoyos gráficos en determinados niveles y para determinados propósitos. Igualmente podrían diseñarse fórmulas híbridas: tras la audición de, por ejemplo, los problemas de una serie de personajes, repartir los consejos más apropiados para cada uno de ellos o, incluso, redactar alguna ampliación de los mismos o algún nuevo consejo para un personaje que carezca de él, bien porque “no hay consejos para todos” bien porque alguno de ellos ha de ser descartado porque no resulta del todo apropiado... Las posibilidades son numerosas, pueden ser creativas, y no hay razón para creer que no sean fiables en su capacidad de medición... alguna muestra de trabajos de este tipo se puede encontrar en determinados DELE.

Hasta ahora hemos revisado los principales modos técnicos de confección de exámenes. Por el camino hemos transitado por algunas destrezas, hemos detectado problemas (bastantes) en términos de administración y fiabilidad, y hemos presentado algunas alternativas (pocas) que pueden proporcionar mayor información al proceso de medición. Soluciones, posiblemente ninguna.

MÁS CONSIDERACIONES GENERALES. *LENGUAS CON FRONTERAS*

En el recorrido hemos observado una nueva constante, profundamente arraigada en el diseño de exámenes pero cada vez más diluida en la práctica docente del aula... Me refiero a la Estricta separación en las destrezas tradicionales que persiste en la arquitectura de los exámenes, con mediciones “exactas” para cada una de ellas...

Y aunque puede ser razonable desde un punto de vista de administración académica e incluso curricular ofrecer informaciones (mediciones / calificaciones) sobre cada una de esas parcelas de conocimiento / competencias de la lengua, lo que no le parece tanto es que la recogida de mediciones se establezca con actividades disociadas de otras destrezas casi exclusivamente (algunas excepciones son tangenciales)...

La observación afecta, desde luego, a lo que hay en las pruebas y a lo que podría o debería haber... Desde luego, la gramática y el léxico están implicados en cualquiera de las restantes pruebas; hemos visto la intervención de la lectura (de textos en sí o de enunciados) en casi todas ellas, con la excepción (en algún nivel) de la interacción oral. Pero, ¿acaso la comprensión auditiva no es inherente a la interacción oral?; ¿no es acaso lo más normal escribir o hablar acerca de lo que se ha oído o leído que responder preguntas cerradas...? Si la lengua como

vehículo de acción y comunicación no aparece segmentada, no parece apropiado presentarla así en una situación de examen. Y no sólo no es apropiado sino que no es tampoco rentable en ningún orden didáctico, desde la relación de los individuos con la lengua hasta el prosaico pero nada banal asunto de la localización de documentos escritos o sonoros, al parecer sumamente específicos para poder desarrollar una explotación que aborde –en el mejor de los casos- una única destreza. Lo “normal” –y económico- en este caso es que un documento presente objetos de aprendizaje (presentación / consolidación / ampliación, y en este caso de medición) variados, y que una explotación sensata, y tanto más en este proceso de medición, recoja esa variedad. Además de datos “eminentemente propios”, esto es, de fácil adjudicación a una determinada destreza en el desarrollo de una explotación, lo común es que existan componentes gramaticales que articulen la significación de un texto escrito, aspectos fónicos que expliquen la interpretación de un enunciado oral, marcadores que en ambos determinen la progresión del discurso, etc. Y desde luego, la posibilidad de articular expresión escrita –de producción libre o semicontrolada- en cada uno de los exámenes.

Abogamos, precisamente por ello, por “EXÁMENES MESTIZOS” EN TÉCNICAS (SIEMPRE LA MÁS PERTINENTE PARA CADA ÍTEM) Y PARA LOS OBJETIVOS DE MEDICIÓN. No se trata de amalgama amorfa, sino de un producto de una gran sutileza lingüística –más adelante veremos algunos ejemplos- en el que se proceda a la medición preferente de una determinada destreza pero que al mismo tiempo incluya ítems adecuados para la medición de otras, y que ésa a su vez reciba informaciones de las restantes, hasta alcanzar un número de ítems suficiente para cada una de ellas... El único “desorden” en todo ello es el que se recaerá sobre los correctores de la prueba, que deberán clasificar y ordenar las diferentes informaciones que contribuirán a confeccionar las mediciones esperadas. Un proceso, en principio muy amanuense, pero no necesariamente, pues en la medida en que hemos definido los exámenes como procesos sin colaboración todo este sistema de administración puede ser fácilmente informatizado e implementado sin apenas inversión. En cualquier caso sería un precio justo: establecer una medición medianamente precisa de la interlengua de un aprendiente en relación con unos objetivos curriculares perfectamente delimitados en la actualidad parece una tarea muy compleja (tanto más si queremos proporcionar retroalimentación), que rebasa la simple superposición de plantillas de corrección sobre hojas de respuestas, pues –no nos engañemos- ése el tratamiento que reciben las denominadas pruebas objetivas.

Esta formulación “mestiza” nos permite no solo una cierta economía en la localización de documentos, sino que DIVERSIFICA LAS POSIBILIDADES DE VERIFICACIÓN DE CONOCIMIENTOS / COMPETENCIAS: un mayor repertorio de muestras de expresión escrita (abierta o semicontrolada, aunque sean fragmentos breves, donde constatar determinados exponentes propios del nivel), nuevos contextos de observación de determinados fenómenos gramaticales o léxicos (de otro modo podríamos llegar a concluir que un alumno que yerre en la elección de *una* forma verbal desconoce *el* sistema verbal español, algo que no es necesariamente cierto; u otro que es incapaz de asociar un sinónimo concreto posee un

vocabulario insuficiente: el diagnóstico certero es que desconoce *ese* término)... Se persigue alcanzar un mayor equilibrio en la medición, que algún modo respete algunas diferencias individuales, en absoluto contempladas en los exámenes, pero previstas en los planes evaluadores.

Este aspecto, el tratamiento de las individualidades o determinados perfiles grupales (nacionalidades, entorno de enseñanza, perfil del docente...) es una de las grandes asignaturas pendientes –y peliaguda– del diseño de exámenes. La alternativa más inmediata es la de INCREMENTAR SENSIBLEMENTE EL NÚMERO DE ÍTEMS Y PROPUESTAS, PERO PRESERVAR EL NÚMERO DE RESPUESTAS DEL CANDIDATO, ESTO ES, OFRECERLE CIERTA CAPACIDAD DE ELECCIÓN. Planteé esta sugerencia hace algún tiempo en la *New Zealand Qualifications Authority*, responsable de los exámenes nacionales que tienen lugar en el país en los tres últimos años de la enseñanza obligatoria, para amortiguar las distancias (sobre todo de vocabulario) entre los candidatos que habían cursado sus estudios con profesores latinoamericanos o no nativos formados allí y los de formación vinculada al español peninsular (entre otras). La propuesta se entendió como “una ventaja” para el alumno (respuesta muy socorrida que delata el tono general penalizador en la percepción de los exámenes) y se descartó. En nuestro entorno pervive en alguna de las pruebas (expresión escrita de los DELE).

Me sigue pareciendo una propuesta justa si el diseño de las pruebas está bien calibrado, y un espacio adecuado para valorar la competencia estratégica del candidato y atender a sus diferencias socioculturales y sociolingüísticas. Un estudiante europeo no abordará la producción oral o escrita acerca de un tema concreto del mismo modo que lo haría un estudiante asiático o africano, pongamos por caso. Tampoco una interacción oral. Y ni siquiera resultará fiable la medición de la competencia gramatical a partir de, por ejemplo, los errores en el uso de los artículos entre estudiantes cuya L1 posea o no sistema de determinación del sustantivo... Si no se prevén estas situaciones el principio ecualizador de los exámenes queda en entredicho, y posiblemente, en lo estrictamente cultural, con una “primacía del primer mundo” (Kumaravadivelu: 2008).

Regresando a lo anterior, incluso podría ser pilotada alguna fórmula progresiva. Un PLANTEAMIENTO BÁSICO: ofrecer un número amplio de ítems con la instrucción de responder necesariamente una cantidad de ellos (elegidos libremente por el estudiante). Un PLANTEAMIENTO PROGRESIVO: la instrucción en este caso exigiría la respuesta de una cantidad determinada de ítems, pero abriría la posibilidad de responder opcionalmente algunos otros de los no contestados inicialmente... El diseño de los ítems debería responder, por una parte, a unos objetivos medianamente estándar y homogéneos, y, por otra, a enunciados basados en el análisis de errores y muestras contrastivas. Un determinado ítem (por ejemplo, alguno relacionado con el artículo) puede ser irrelevante para un estudiante italiano, pero sumamente significativo para un aprendiente ruso; la inclusión de repertorios opcionales podría cubrir esos espectros, esto es, un conjunto de ítems que se añadirían a la propuesta común, especialmente concebidos en razón de las dificultades que puedan entrañar en relación con la L1 del

aprendiente, y cuya cuantificación resulte variable, en función precisamente de esa L1.

La gestión no es fácil, aunque se puede simplificar con un tratamiento informático, que contemple un etiquetado de todos los ítems. La información recibida, extremadamente poderosa: qué ítems ha abordado el candidato, cuáles no, el índice de corrección, y la relación de todo ello con, por ejemplo, la nacionalidad o la L1 del estudiante, la situación o no de inmersión, su trayectoria de aprendizaje, su biografía lingüística, etc. (esto es, informaciones neutras o pseudoanónimas, no necesariamente comprometedoras)... Posiblemente no sería posible adoptar decisiones inmediatas en lo referido a la calificación del examen, pero sí respecto a la evaluación en general y al diseño mismo de pruebas de medición. En cualquier caso, los diseños de pruebas con este perfil son complejos y se deberían asentar sobre investigaciones y estudios de campo relevantes.

Hemos apuntado arriba alguna referencia a aspectos socioculturales. La práctica totalidad de los desarrollos curriculares (desde el *MCER* al *PCIC*, y desde luego los documentos de las EEOOI) contemplan de modo específico esos contenidos, pero sin embargo no se articula ninguna forma concreta (a lo sumo alguna recomendación) de evaluar / medir esa competencia. Las obvias alusiones a material real se fundan en lo que el *PCIC* denomina “Hechos y referentes culturales”, con frecuencia distantes para algunos candidatos y comúnmente efímeros.

Dejando de la complejidad –y tal vez pertinencia y necesidad de medición, aunque ése es otro asunto- de determinados contenidos curriculares (en especial los que aluden a “Hechos y referentes culturales”⁷), lo cierto es que las variables culturales no aparecen reflejadas en la medición de la competencia, y desde luego es posible rastrearla en muchas de las producciones de los aprendientes.

⁷ Sobre este aspecto, vinculado al conocimiento del mundo, viene al caso una anécdota referida por mis compañeros de Departamento. En el examen de interacción oral (nivel A1) un estudiante europeo y otro chino debían organizar un viaje de fin de semana. La propuesta del estudiante europeo fue visitar París; el estudiante chino se mostró de acuerdo. El estudiante europeo, mejor conocedor de las dinámicas de un examen de lenguas y capaz de desplegar estrategias en este sentido, fue sugiriendo posibles lugares de visita (la Torre Eiffel, el Museo de Louvre...) siempre con el acuerdo de su compañero asiático. La última de las sugerencias –un paseo por los Campos Elíseos- obtuvo en cambio un rechazo tajante: *Es que a mí no me gusta el campo...*

porque siempre hay alguna sorpresa me sale de
 una tienda pequeña, que increíble.
 Barrio Gótico es el corazón de Barcelona, es
 un buen lugar para visitar.
 Ahora te enseño una construcción, es la Catedral
 de este barrio, se empezó a construir en año
 1298 y se terminó en el año 1454. Dentro de
 Catedral hay una fuente, si pongas una moneda
 y hagas un voto o una promesa a un buda,
 seguro que tendrás las cosas que quieres....
 Otras construcciones, es ^{SOR} Murallas, pero no es la
 muralla de China, en Barcelona hay unas ~~murallas~~
 murallas que se llaman Murallas Romanas que
 median nueve metros de alto por tres y medio
 de ancho y 1270 metros de longitud.
 En este barrio también hay otras construcciones que
 no podemos ver en China, nunca has visto unos
 edificios de forma no parecen un edificio ^{¿No?}, pero
 en Barcelona si que puedes.
 Bueno, pues eso es todo, Si vendrás a Barcelona
 te enseñare ~~todo~~ todo lo que he visto y que Barcelona
 te guste mucho.
 Un besito!
 Linguiiao.

En su texto, esta estudiante china (B1) deja constancia de al menos tres manifestaciones socioculturales: la catalogación de realidades nuevas o desconocidas ("la catedral de ESTE barrio"), la interpretación de la realidad cultural ("una promesa a UN BUDA") y la comparación con su C1 ("no es la muralla de China", "otros construcciones que no podemos ver en China"), datos todos ellos que no encuentran espacios en los exámenes para generar un diagnóstico complejo del candidato.

Por poner algún ejemplo (Escandell: 1996 / 2009), la expresión de peticiones en español se puede realizar con imperativos, preguntas indirectas y sugerencias. Articular un ítem basado en, pongamos por caso, en una pregunta indirecta representará un esfuerzo añadido para un estudiante chino, en cuya L1 la articulación de peticiones no contempla las estructuras interrogativas, y lo será también, aunque en distinto grado, para un estudiante ruso, capaz de reconocer la interrogativa como forma de petición a pesar de que no la use como tal. Si la elección es el imperativo (quizá la más habitual en español), el que antepondrá sus filtros culturales y sus interferencias sociopragmáticas podría el estudiante anglosajón, quien preferirá otras formas indirectas para expresar la petición. Lo mismo puede suceder con un estudiante alemán que deba articular consejos o

responder algún ítem relacionado con ellos, si no posee la condición de “experto” que le legitime para hacerlo... De la misma manera un estudiante chino tenderá a no contradecir en ningún momento a su interlocutor, de modo que un examen que incluya una instrucción orientada a debatir y ponerse de acuerdo puede verse notablemente perturbada en función de las intervenciones del interlocutor. En la misma tarea relatada en la nota 6, en este caso entre dos estudiantes chinos, la propuesta de organizar una salida de semana se vio casi frustrada: a uno de los participantes le era imposible el sábado; al otro, el domingo; su incapacidad para la simulación y para oponerse al interlocutor obstaculizaron el cumplimiento de la tarea en esa situación de examen. La misma actividad funcionó sin problemas destacables con estudiantes de otras lenguas, con una identidad cultural próxima, compartida o asumible.

Es desde luego del todo imposible prever toda esa casuística en estas situaciones de examen, pero tampoco se puede ser ajeno a ella. La competencia en una LE, y en este caso la medición, no se puede hacer sobre una eliminación del *background* cultural de los candidatos, exigiendo una metamorfosis cultural (Kumaravadivelu, 2008). De hecho los candidatos de las pruebas de francés o alemán, o incluso chino, completaron la referida tarea (todos hemos experimentado situaciones similares como aprendientes de lenguas), pero ¿lo hicieron como lo hubiera hecho un francés, un alemán o un chino nativos?

Es aquí donde nuevamente cobra sentido, la diversificación de ítems y tareas en los exámenes, como una medida reparadora –solo en parte- de esas situaciones. La capacidad del candidato para elegir, sin que se resienta la capacidad de medición por parte del examinador.

CONSIDERACIONES PARTICULARES. LAS PRUEBAS

Abordamos a continuación el análisis de algunas de las pruebas, que hemos identificado como más o menos constantes. No es un análisis exhaustivo y la ejemplificación, por razones de espacio, ha de ser necesariamente sucinta. Todos los docentes tenemos algún grado de implicación con los exámenes: los apuntes siguientes son simplemente puertas abiertas para facilitar la entrada de nuevas ideas (que debemos construir entre todos) y, ya de paso, puesto que están abiertas, expulsar alguna que otra que se nos haya quedado vieja y rancia.

Cabe, en primer lugar, determinar QUÉ CONTENIDOS QUEREMOS MEDIR Y CUÁL ES EL MODO MÁS EFECTIVO DE HACERLO, de entre las técnicas que tenemos a nuestra disposición, incluidos instrumentos híbridos. Como es natural, los contenidos los establece el marco curricular de nuestra institución docente, pero vaya por delante que no es posible desarrollar tareas de comprobación para todos ellos. Y eso es algo que no siempre se ha comprendido bien. Entendemos por tanto que ES PREFERIBLE LA REAPARICIÓN DE ÍTEMS EN FORMAS PRÓXIMAS QUE CONFIRMEN LA ADQUISICIÓN DE UN DETERMINADO CONTENIDO QUE NO EL COQUETEO CON UNA GRAN DIVERSIDAD DE ELLOS (recordemos que, según lo que hemos visto, el grado de

fiabilidad de las técnicas no es excepcional). Además el añadido de indicadores de validación puede resultar consistente. Existen pruebas, con frecuencia las de interacción oral, en las que se pide al candidato que describa una fotografía, que opine sobre lo que representa y que relate alguna experiencia vinculada a la imagen en cuestión... O sea, la voluntad de querer constatar todo.

En relación con ello, un concepto claro del reparto entre las distintas pruebas... Sirva un ejemplo. A partir de A2 y en B1, el relato, la narración, se convierte en objeto casi obsesivo. Es cierto que en él se dan cita muchos de los contenidos (preferentemente gramaticales) "más objetivables" de las competencias esperadas... Sin embargo, en un reciente trabajo de preparación de exámenes "internos", y tras un proceso de elaboración ciertamente no demasiado coordinado nos dimos cuenta de que todos y cada uno de los documentos y propuestas productivas giraban en torno a la narración. También nos dimos cuenta de que la inclusión de actividades productivas vinculadas al relato en A2 podrían mostrarse obstaculizadoras de la tarea, por la casi necesidad de aparición del estilo indirecto, todavía no presentado en ese nivel, de modo que era aconsejable reconducir esos contenidos hacia otro tipo de trabajo o bien con instrucciones circunscritas más a la narración que al relato de anécdotas en pasado, sin referencias, pues, a las experiencias personales (de las que tiramos a menudo).

Aunque los ejemplos pueden parecer obvios, la puesta en práctica no lo es. Exige coordinación y también un trabajo notable de búsqueda y selección de documentos capaces de cubrir todos los contenidos objeto de medición y de brindar oportunidades de reaparición validadora para cada uno o bastantes de ellos. Y, desafortunadamente en estos casos, el procedimiento suele ser el inverso: localizamos un "buen texto" y de él extraemos los ítems que dé de sí (Lozano y Ruiz Campillo: 1996 / 2009). Las desventajas de este procedimiento pueden ser aminoradas si formulamos explotaciones híbridas: ese "texto hallado" puede ser fuente para ítems variados que, junto con los restantes de las demás pruebas, pueden completar ese espectro variado y diverso que perseguimos (eso mismo es muy difícil encontrarlo en un único texto, ajustado a un nivel, con una extensión razonable y una cierta cohesión informativa).

La selección y determinación de los documentos han determinar el tipo de prueba (objetiva / subjetiva, o reaparición en ambas) y la ordenación de las partes de la misma, con arreglo a la carga que comporten los documentos y los ítems y las tareas, y también los objetivos que hayamos establecido. Por ejemplo, si una lectura dispone de un repertorio "útil" de marcadores temporales, pongamos por caso, no parece inadecuado establecer a continuación una prueba de expresión escrita en la que el candidato tenga oportunidad de producir parte de ese repertorio en su propio escrito... Algunos casi denominan eso "copia", pero seguramente es sobre todo una excelente oportunidad de aprendizaje (y de medición) en la que el aprendiente es capaz de demostrar un buen trabajo de comprensión lectora en lo que se refiere a la ordenación textual que le es útil para una producción en un nuevo contexto... (y tal vez por ello nos podría

permitir cuantificar su competencia en, al menos, dos categorías / destrezas diferentes).

Un último apunte: LAS INSTRUCCIONES DE LOS EXÁMENES. Es obvio: han de ser extremadamente claras (en ocasiones, en los niveles iniciales, la lengua empleada rebasa la competencia esperada), y lo han de estar preferentemente antes de la administración de la prueba (si es preciso, con informaciones en la(s) L1 de los candidatos, y de ser posible, con las oportunas comprobaciones de que realmente han sido entendidas). Por otra parte, es destacable que pruebas como los DELE o los certificados de las EEOOI se sirven de un tratamiento de usted; aunque sea la forma usada en muchas aulas americanas, lo cierto es que, al menos en España y los profesores formados en España, utilizan mayoritariamente el tuteo en el habla de la clase, lo que da lugar a una colisión entre esa lengua de aprendizaje y la lengua del examen, con posibles implicaciones, además, en la afectividad de la situación.

COMPRENSIÓN ESCRITA

Articula la práctica totalidad de las pruebas, pero solo es cuantificada en la sección dedicada a ella. Como hemos apuntado y descrito, es casi un callejón sin salida: son las preguntas las que tienen “trampa” en opinión de los estudiantes, no los textos, y, sin duda, a veces bloquean el acceso a la información original con redactados más o menos crípticos que todos hemos padecido como estudiantes de lenguas y que aun en los exámenes que administramos hemos de releer en alguna ocasión.

Es por lo general una destreza sólida, la que suele inspirar mayor confianza a los aprendientes, y en la que despliegan con mayor efectividad sus recursos estratégicos, tanto en el aula como en situación de examen. Por esa razón, es seguramente EL ENTORNO EN EL QUE SE PUEDEN ASENTAR LOS COMPONENTES DE MAYOR “RIESGO” DE CUANTOS SE QUIERAN VALIDAR, entre ellos los contenidos gramaticales, por lo demás siempre amparados en textos escritos, aunque en esos casos –las pruebas exclusivas de gramática- no promuevan estrictos ítems de comprensión escrita: los recursos movilizados son otros.

No obstante, y a pesar de ello, el tipo de ítems propuestos en muchos de los exámenes es pobre, e incide especialmente sobre el vocabulario y, a lo sumo, su cotexto más inmediato. Los enunciados de las preguntas se articulan sobre sinónimos y paráfrasis que han de ser desvelados en el texto original. LA MEDICIÓN SE SUELE ESTABLECER SIN APOYO DE ÍTEMS BASADOS EN FENÓMENOS DE INTERPRETACIÓN, a pesar de que es enteramente posible.

15. En relación con la actitud de Carter sobre la fotografía del buitre y la niña, el autor del texto

- a) no muestra comprensión alguna por Kevin Carter
- b) critica a Kevin Carter
- c) parece entender la actitud de Kevin Carter



Ejemplo de pregunta de comprensión e interpretación general empleado en un examen de comprensión escrita (B1+). En este caso, por la complejidad y extensión del texto, se emplazó en último lugar, pero es razonable incluir algún ítem general, de interpretación, al comienzo para procurar la lectura completa. Posiblemente una de las formas más amables –que puede descargar de tensión al aprendiente, en especial, si ésta es la primera prueba del examen, como sucede en los DELE- sea el ya citado procedimiento de incluir una fotografía que debe ser vinculada con el texto en algún modo. Los grados de abstracción de la misma se han de asociar con el nivel.

Por otro lado, LOS ÍTEMS SE PRESENTAN CASI SIEMPRE EN ORDEN SECUENCIAL: la respuesta a la primera pregunta se sitúa en las primeras líneas de texto; la segunda, en las posteriores, y así sucesivamente. El planteamiento genera unas inercias negativas⁸ en el trabajo de comprensión: que los candidatos acostumbren a desarrollar estrategias (certeras tal y como están configuradas las pruebas, pero pedagógicamente inadecuadas) según las cuales comienzan a contestar sin haber completado una lectura total del texto, y así pueden dar por válida una información que se desmiente en el párrafo siguiente, por ejemplo; aunque ciertamente los ítems no suelen reparar en ello: se valida exclusivamente la literalidad del enunciado, lo que da pie a que la unidad de comprensión no rebase una porción breve de texto, y en consecuencia diluye la atención en los componentes macrotextuales y en la progresión de la información. Los DELE mantienen un trabajo de asociación entre textos -la versión d) de los enumerados arriba, con las complicaciones apuntadas entonces-, si bien en la mayoría de las muestras observadas, la técnica es utilizada para asociar preguntas y respuestas de una entrevista: por lo general y según hemos visto arriba, la actividad se resuelve en las primeras palabras de las respuestas, de modo que en ocasiones no es necesario completar la lectura este tipo de ítems, con esta formulación concreta. Es cierto que es ésa una estrategia frecuente en la interacción con un texto, pero en situación de examen se desaprovecha un documento que nos podría haber aportado otras informaciones de medición.

Si nos distanciamos de esa tendencia, además de ítems relacionados con la interpretación, nos pueden ser útiles como validadores del trabajo de comprensión lectora determinados ÍTEMS INTRATEXTUALES, ESTO ES, QUE IMPLIQUEN EL MANEJO DE INFORMACIÓN ENTRE DOS O MÁS PÁRRAFOS, intercalados entre los restantes ítems (para no ahondar en la disposición secuencial) y que incidan, por ejemplo, en la progresión de la información (comparación de informaciones, evolución de un suceso, datos nuevos y su relación con las informaciones previas, actitud del autor...).

⁸ Tras una lectura u otro proceso de recepción, lo habitual en los intercambios humanos es preguntar por lo más interesante, lo más sorprendente de la información adquirida, no sobre lo que aparece en primer lugar, en segundo lugar, etc.

Igualmente entendemos que LA MUESTRA MÁS FEHACIENTE DE LA COMPRESIÓN DE UN TEXTO ES EL RESUMEN / PARAFRASEO que pueda hacerse de él y la construcción de nuevo discurso a partir del mismo. La dinámica de un examen, al menos en formato tradicional, impide un trabajo oral, de modo que la opción de escritura abierta puede ser apropiada: se trataría de un texto breve, sin mayores riesgos. Para evitar la tentación de la copia literal se puede arbitrar alguna solución intermedia (véase abajo, bajo el epígrafe de expresión escrita). En cualquier caso, la PRESENCIA DE AL MENOS UN ÍTEM DE REDACCIÓN ABIERTA NOS PARECE INELUDIBLE EN CUALQUIER TRABAJO RELACIONADO CON LA COMPRESIÓN ESCRITA.

Por otra parte, junto con los ítems “argumentales” y los citados de interpretación, la medición de lo gramatical tiene perfecta cabida, aunque únicamente en la medida en que se asiente en la significación, tal y como se sugiere en los ejemplos...

11. LÍNEA 27

pero pensar que tantas historias se iban a perder... (11)

Señala la opción que indica el significado de SE IBAN A PERDER:

- ☐ a) las historias ya se habían perdido; por eso el cómic las presenta
- ☐ b) las historias aún no se habían perdido
- ☐ c) algunas historias se han perdido pero otras todavía no

Ejemplo de ítem esencialmente gramatical, asentado en el significado del texto, presentado con un grupo de B1+. Por otra parte, el enunciado de la pregunta intenta ser gráficamente autosuficiente, con la opción de respuesta en la misma página.

10. A Carter, la pregunta “Y después, ¿ayudaste a la niña?” (en el párrafo situado entre las líneas 72 y 74) se la hacía:

- a) nadie en verdad; era una hipótesis, pero se la podía hacer mucha gente
- b) solamente su familia, pero muy a menudo
- c) en todos los lugares

☐

11. Después de publicar su fotografía de la niña y el buitre, a Carter le decían

- a) que si habría ayudado a la niña
- b) que si había ayudado a la niña
- c) que ayude a la niña

☐

Otra muestra de intervención gramatical. La pregunta 10 actúa en cierto modo como capacitadora de la siguiente, en la que el candidato ha de reconocer el desarrollo morfológico del estilo indirecto, pero sin perder de vista el significado.

A las sugerencias apuntadas, hemos de añadir que la localización de textos no es extremadamente demasiado difícil, en especial en la internet, con algunas palabras clave, y que en cualquier caso, es posible adaptar cualquier texto para que se ajuste a los cometidos de medición que queremos asignarle. No obstante, una vez seleccionado el texto (en principio auténtico, entendido como cualquier texto que no se dirige explícitamente a un lector extranjero), es razonable preservar ciertas características gráficas (Yagüe: 2008): formato aproximado (noticia, anuncio, carta...: ¡hay vida más allá de la *Times* pto. 12!), legibilidad (selección de fuentes apropiadas, y portadoras de afectividad), preferentemente distribución del texto en una única cara, numeración de las líneas, interlineado o

márgenes generosos (para posibles anotaciones, si la institución administradora del examen las autoriza), eventual inclusión de fotografías o imágenes (ilustrativas o informativas, con posibles trabajos diferentes en cada caso), etc. Algunos de estos factores contribuyen a generar inferencias y ajustar conocimientos previos; otros, a facilitar el desarrollo “técnico” de la prueba. Todos son necesarios.

Sobre la selección argumental, imposible decir nada que no resulte demasiado etéreo: los conceptos de texto interesante, ameno, divertido, que suscite reacciones u opiniones, y similares son vagos para candidatos sumamente diversos. Lo que es cierto es que un hablante de cualquier lengua no lee, o deja de leer, un texto que no cumpla esas expectativas. Si la prueba se articula sobre un par de textos breves o de extensión media, se puede proponer un tercero para que el candidato elija dos de esos tres (lo que representa una explotación añadida, y que cualquiera que sea la elección permita acceder a los contenidos objeto de medición).

COMPRENSIÓN ORAL

“Cuando uno escucha muy atentamente no ve nada,
como si cada sentido forzado al máximo excluyera el ejercicio de los otros”
Javier Marías: *Corazón tan blanco* (Random House Mondadori, pág. 50)

Se articula como una medición incomprensiblemente sesgada (Pastor: 2009). ÚNICAMENTE SE PLANEA PARA MENSAJES EN QUE EL RECEPTOR NO TIENE CAPACIDAD DE INTERVENIR, una parte –solo una parte- de nuestra actividad como oyentes. Obviamente se ha de dar cabida a pautas de medición bien establecidas en la interacción oral, la que tiene lugar entre iguales, con otro(s) aprendientes, y con hablantes nativos o competentes (examinadores o no). Ése es seguramente su entorno natural. Y también en él se podría incluir, como tarea, algún documento sonoro (común o individualizado, en este caso podría ser escuchado en un reproductor individual tipo iPod proporcionado por los examinadores) con la instrucción de que el candidato transmita verbalmente esa información a los compañeros de la prueba). No es una gestión difícil, tiene un propósito comunicativo definido y la(s) instrucciones pueden ser muy básicas, orientadas a algo tan frecuente en el habla como explicar verbalmente algo que hemos oído.

De gestión algo más difícil pero con posibilidades de acomodación en soporte informático, como sugeriré más abajo, la posibilidad de incluir algunos ítems no como enunciados escritos sino auditivos en el resto de las pruebas. El concepto de esos enunciados no sería la simple lectura del enunciado escrito, sino un mensaje algo más elaborado (lectura de un pasaje y vincularlo con otro de significado semejante en el texto, aportación de informaciones para localizar la que se encuentra o no en el texto, alguna discriminación fónica que no ha sido imposible encontrar en documentos auditivos o audiovisuales). De nuevo, la medición de LA COMPETENCIA AUDITIVA DEL CANDIDATO PROVENDRÍA DE DIVERSAS FUENTES,

DE DIVERSOS CONTEXTOS Y DE MAYOR NÚMERO DE HABLANTES: la capacidad de proporcionar información devendría probablemente más rigurosa si el diseño de la prueba responde a unos parámetros de calidad.

En su formulación actual, las pruebas de comprensión auditiva se reparten en las que tienen soporte auditivo exclusivamente y las que se ofrecen en presentación audiovisual, éstas últimas ganando terreno a las primeras, a pesar de que se limita el número de documentos disponibles y de que la manipulación y la edición (la segmentación y los ajustes de duración y la preservación de la calidad visual) son sensiblemente más complejas para los no profesionales.

En el caso de LA PRESENTACIÓN AUDIOVISUAL, EL ESFUERZO COGNITIVO EXIGIDO A LOS CANDIDATOS ES EXTREMO. El visionado, la audición, la lectura y la escritura actúan de forma más o menos simultánea para intentar adentrarse en un tipo de mensaje efímero por naturaleza: la respuesta a los enunciados ha de ser simultánea o basarse en la memoria auditiva reciente. Complejo asunto, que puede comprometer el concepto de viabilidad (*MCER*).

Además, parece detectarse un tratamiento viejo en moldes nuevos. Es decir, a falta de estudios que sigan legitimando⁹ la intuición, la información visual, poderosísima sin duda, y así ratificada por los expertos, carece de explotación en los ítems de estas pruebas, que siguen subrayando de hecho que la comprensión auditiva es comprensión auditiva y nada más. Con tales presupuestos las imágenes no son, en efecto, necesarias, ni siquiera pertinentes. De hecho, y a pesar de no que no he hallado estadísticas rigurosas al respecto, si observamos a nuestros estudiantes durante una prueba de estas características comprobaremos que la atención a la pantalla es mínima: bien se concentran en el papel bien cierran los ojos para incrementar su concentración, tal y como escribe Javier Marías en la cita que abre este apartado. Y en algunos casos, la situación deriva en desastre, especialmente en escenas o reportajes en los que interviene más de un hablante (a veces identificado por algún texto en pantalla). Puede ocurrir que los redactores de pruebas den por supuesto que el candidato ha estado observando atentamente todo el vídeo y, así, funden el redactado del enunciado en algún referente visual (*Según fulano de tal..., La persona que está en...*) para a continuación preguntar sobre el mensaje estrictamente oral. Despojado de la información inicial, el estudiante encuentra problemas ajenos a la comprensión misma, ya que se topa con los referentes a los que alude el enunciado.

LAS EXPLOTACIONES MÁS COMUNES VUELVEN A INSISTIR EN LA CARGA LÉXICA: los enunciados

⁹ He tenido la oportunidad de dirigir una interesante memoria de máster (Instituto Cervantes – UIMP) sobre este asunto, con resultados francamente reveladores (a pesar de que el número de casos estudiado era bajo). Se administró una misma prueba a diversos grupos, a unos de ellos en formato audiovisual y a otros exclusivamente la audición, sin imágenes. Las explotaciones previstas para ambos casos eran “preferentemente auditivas”. El segundo de los grupos obtuvo mejor rendimiento (calificación que el primero). Los informantes del primero apuntaron en el cuestionario preparado que lo visual no había ayudado o había entorpecido la resolución de la tarea (Casañ, 2009).

se redactan de nuevo con sinónimos y paráfrasis, con el agravante ahora de que el tiempo de reacción es mucho más limitado que en la comprensión lectora. A ellos se les acostumbra a añadir alguna pregunta adscrita a discriminación fónica. Y en ambos casos suelen corresponder enunciados de tipo a), b) o d), los más extensos, que comprometen el desarrollo de la prueba. En alguna ocasión he cronometrado el tiempo invertido –con mi “competencia” de hablante nativo– en leer con antelación los enunciados de la preguntas, sin procurar una comprensión minuciosa, y en la totalidad de los casos el tiempo invertido ha sido superior al que ofrecía la prueba.

Por fin, una gran ausente: la habitual falta de ítems relativos a prosodia y, en el caso del soporte audiovisual, la gestualidad, expresiones faciales, etc., claves todas para determinar LA INTENCIÓN COMUNICATIVA DEL HABLANTE, DESATENDIDA IGUALMENTE EN LAS PRUEBAS, QUE SIGUEN REAFIRMANDO LA LITERALIDAD DEL MENSAJE LINGÜÍSTICO. En este sentido, recientemente ensayamos con un ítem de cierto riesgo, que incidía en estos aspectos. En un documento monologado, una mujer relataba recuerdos de su juventud; el enunciado de la pregunta pedía que el estudiante dictaminara si era verdadero o falso lo siguiente: “Según lo que ella explica, a los padres de la persona 1 no les parecía correcto que acompañaran a su hija a casa”. La producción literal de la mujer no contenía constituyentes lingüísticos determinantes para responder a la pregunta, pero sí extralingüísticos (...) *Siempre me acompañaba uno..., y eso... mis padres, claro, lo que pasa.*

El panorama, difícil, se completa con la necesaria atención a diversos planos visuales (la pantalla y el papel o papeles). Las aportaciones de la psicología cognitiva (Mayer: 1997, 2002) acreditan que esa duplicidad de planos visuales resta concentración al estudiante. La ubicación en un mismo plano del documento audiovisual y de la hoja de trabajo pueden aliviar la situación: el escenario no puede ser otro que la pantalla de un ordenador.

Pero todas estas peculiaridades de la prueba de comprensión auditiva no nos deben hacer olvidar la esencia comunicativa de este tipo de acciones: uno escucha para anotar un teléfono, los horarios de los trenes para volver de la playa, la dirección de un restaurante, y escribe, producciones algo más extensas aunque no siempre con cohesión textual, las ideas principales de una clase o una conferencia (si bien este discurso tiene sus propias peculiaridades, en dicción, ordenación y redundancia)... Casi ninguna de esas acciones (excepto en niveles iniciales) está contemplada en el diseño de exámenes de medición de la comprensión oral. No se trata de desembarazarse necesariamente de las existentes, pero sí de dar cabida a las más inherentes a la comunicación humana.

2. Presta atención a lo que dicen estas personas y explícalo con tus propias palabras.



--

Ejemplo de producción abierta tras la audición (B1+). Las mujeres del vídeo explicaban tres informaciones relevantes (dato clave para la cuantificación) y algunos detalles y experiencias vinculadas a esas informaciones. La extensión del texto del candidato se limitaba voluntariamente a la extensión del cuadro. Eventualmente, la respuesta podía aportar información sobre el grado de dominio en el uso del estilo indirecto (aunque su empleo no era necesario aquí).

Las formas de amortiguar ese efecto indeseado del vídeo no están demasiado estudiadas, o al menos no está comprobada su efectividad. Podría ser sensato ofrecer una visualización del documento sin ningún cuestionario asociado a fin de procurar una comprensión global, para a continuación proporcionar el cuestionario. El examen requeriría entonces tres visionados (el previo, y los dos habituales, y a con hojas de trabajo), lo que podría ser considerado una "ventaja", y desde luego una situación inédita en la vida real, donde a excepción de algunos anuncios públicos, los mensajes no se repiten, y donde la petición de aclaración puede dar lugar a una repetición (seguramente una reformulación), pero no en tres ocasiones. No obstante, algunas de las muestras del DELE A1 incluyen tres audiciones, cuando se trata de un discurso medianamente elaborado (únicamente dos, cuando el documento oral se funda en enunciados breves, o en los restantes niveles). Todo lo anterior no debería anular la posibilidad de PROPONER UNA PRUEBA FUNDADA EN UNA ÚNICA AUDICIÓN, de carácter breve y con unos ítems bien determinados que recuperasen aspectos generales del contenido. Este tipo de prueba, en contraste con las restantes, fundadas en dos audiciones, puede resultar significativa (proporcionar una valiosa información complementaria) a efectos de medición de la competencia auditiva: así sucede en la vida real con mensajes radiofónicos, televisivos o similares en los que el receptor no puede interactuar con el emisor. Para preservar esos rasgos, el encaje de esa prueba debería ser en la parte específica, ya que en otros momentos (la interacción oral) podría tener cabida de forma natural la solicitud de repetición o aclaración.

Alguna de las opciones de las que me he servido puede ser útil, aunque no he podido verificar su validez. Se fundan sobre todo en CAPTURAS DE PANTALLA, que se insertan en la hoja de preguntas y que pretenden recuperar la memoria visual (previendo que el visionado ha sido incompleto) y asociarla con la memoria auditiva.

Otro de los objetivos son preguntas abiertas, de comprensión global, o incluso de creación e interpretación, en consonancia con lo que explicaba arriba, como la del ejemplo.

PRIMERA AUDICIÓN

Vas a escuchar a dos personas que hablan de sus recuerdos. Contesta las preguntas de la manera que se pide en los enunciados...



PERSONA 1



PERSONA 2

PERSONA 1 (2 minutos)

2. De acuerdo con lo que dice, usa tres adjetivos para describir el *CARÁCTER* de persona número 1. Explica por qué has elegido esos adjetivos.

Considerando que el léxico complejo de descripción física y de carácter formaba parte del temario de curso (B1+), uno de los primeros ítems pedía a los candidatos que escribieran una serie de adjetivos para describir a las dos personas que hablaban, y que en una línea justificaran la elección de esos adjetivos. En una sesión posterior (la última del curso), en plenaria, pactamos los adjetivos aceptables: el repertorio no fue especialmente extenso, y los estudiantes consideraron la pregunta de "dificultad media".

La elección de este tipo de preguntas emplazadas al inicio de la prueba responde además a la voluntad de DISMINUIR LA ANSIEDAD. Se trata de la prueba sobre la que el aprendiente dispone de menos control, y no es infrecuente el bloqueo para la totalidad de la audición. Se quiere que la(s) pregunta(s) iniciales generen confianza y no se persiga un dato extremadamente específico que tal vez escape a la comprensión del candidato. Es un aspecto que por lo general no se tiene en cuenta en la redacción de ítems (la explotación da de sí en función de las características del material sonoro, y sí el primer ítem es complejo, pues qué le vamos a hacer). Además el aspecto de los enunciados no exige una comprensión ardua: la instrucción de lo que se debe hacer es clara, y el redactado de la respuesta se completa con posterioridad.

Respecto a ítems que prevén la redacción de respuesta por parte del candidato, alguno de ellos aparecen en el DELE A1, que elimina así la corrección automatizada. La observación de algunas de esas propuestas pone de manifiesto que la actividad se funda exclusivamente en el reconocimiento fónico de unas serie de piezas léxicas: sólo se admite una única respuesta (la realmente pronunciada).

Teresa es una mujer que busca pareja y llama a la radio para hacer una descripción personal. Complete el texto con la información que falta. Escuchará la audición tres veces.

Escriba las palabras en la Hoja de respuestas.



Buscando pareja

0. Teresa es una mujer joven y... trabajadora.

19. Teresa está soltera y vive con su
20. A ella le gusta viajar y leer.
21. Es de literatura.
22. A Teresa no le gusta la
23. Su casa es muy
24. Los tiene clase de baile.
25. Quiere conocer a un chico

Tarea del DELE A1 de mayo de 2009 que exige el redactado de las respuestas. Sólo se admite una única respuesta, de modo que el trabajo está más próximo del reconocimiento léxico que del de comprensión auditiva, sin embargo los actividades de vocabulario no incluyen inputs auditivos. Por lo demás la ilustración ofrecida no resulta significativa para la resolución de la tarea. El audio, de perfil inconsistente en muchos aspectos, está disponible en http://diplomas.cervantes.es/docs/ficheros/200906180001_7_3.mp3

No conozco estudios de resultados aplicados al ELE que determinen las ventajas o desventajas de una audición larga respecto a una audición corta. La primera genera cansancio, pero al mismo tiempo –si es compacta- puede proporcionar datos al oyente para recuperar informaciones previas que le hayan sido requeridas, algo que no acostumbra a suceder en las grabaciones breves. Parece a priori sensato proponer ambos tipos de audiciones y, en todos los casos, informar al candidato de la duración del episodio que oirá, para que pueda calibrar sus estrategias de comprensión, en muchos casos fundadas en la reconstrucción de segmentos aislados (como de algún modo representa el dibujo siguiente). Y eso nos subraya de nuevo el gran esfuerzo cognitivo que representa una explotación orientada también a segmentos aislados.



La capacidad para elegir responder dos audiciones de tres, por ejemplo, está muy comprometida en la actualidad por el carácter plenario de la audición, pero la audición individualizada puede ser una realidad de fácil ejecución con un soporte informático y los correspondientes cascos, donde, como, hemos dicho, convivirían en un mismo plano visual los elementos de la audición. Ello facilitaría también la regulación del volumen de audición, aspecto comúnmente no contemplado por los administradores de pruebas, si bien diversos estudios (Pastor: 2009) certifican que los apren-

dientes precisan de un volumen algo superior al que demandaría un hablante nativo.

Respecto a los modelos de texto, parece adecuado extender la medición a dos géneros básicos: el monológico y el conversacional, con requerimientos específicos cada uno de ellos, y en algunos casos dificultades añadidas insuficientemente previstas: la comprensión de algunos documentos sonoros se complica por la carencia de conocimientos sobre el tema. La comprensión puede ser posible, pero el resultado acostumbra a ser sesgado y mecánico...

En cuanto al carácter de la audición, en las EEOOI de Cataluña siempre hemos preferido documentos audiovisuales auténticos; los DELES, en cambio, se presentan como grabaciones, en ocasiones como un extraño híbrido: se mantienen (con adaptaciones) las características de contenido de un texto real, pero no la dicción, ya que el texto es leído por un locutor, incluso con la paradoja de que documentos procedentes de archivos sonoros latinoamericanos¹⁰ y así referenciados (por ejemplo, de la radio) son emitidos con una variedad de español peninsular. La falta de naturalidad es especialmente marcada en muchos de los textos dialogados. Ignoro las razones por las que se adopta esta configuración, quizá por problemas derivados de los derechos de autor. En las EEOOI de Cataluña entendemos que es improcedente un texto no auténtico – también en su dicción – para unos estudiantes que viven en un contexto de inmersión, si bien esta decisión entraña innegables dificultades técnicas y de explotación. La inclusión de variedades del español es un asunto controvertido (en relación con políticas de difusión de la lengua, pero también de contacto con ella: origen de los materiales docentes, procedencia del profesor).

En cualquier caso, por el momento, resulta difícil y arriesgado apostar por un peso importante de lo visual si no tenemos ni la certeza (ni la tradición) de que los estudiantes miren el vídeo. En este sentido, sí he ensayado como actividad de clase (A2 y B1+) el uso de vídeos en el trabajo de comprensión lectora: los vacíos

¹⁰ No siempre aparece identificada la procedencia de todos los documentos sonoros que se proponen en los DELE. Se infiere que los que no se identifican son textos creados para la ocasión.

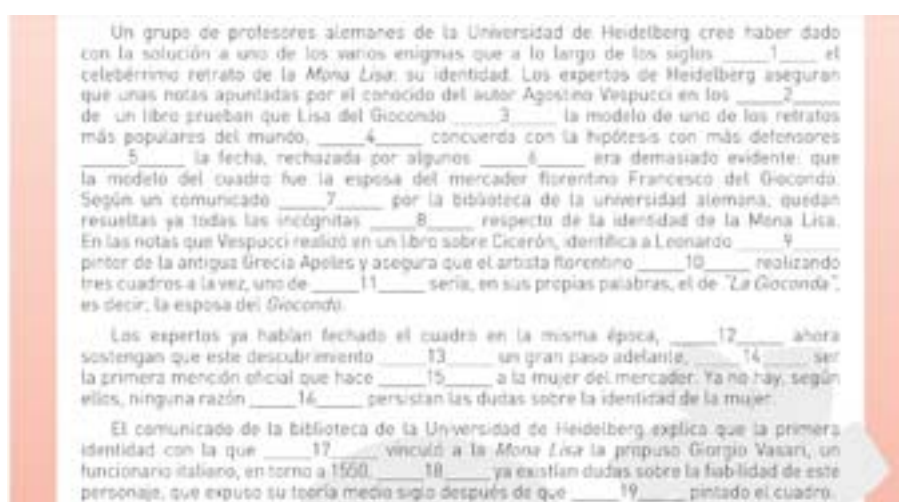
de información del texto escrito –que era necesario comprender- se resolvían en parejas con el visionado de un vídeo, con o sin audio. De nuevo el mestizaje.

GRAMÁTICA (Y VOCABULARIO)

Prueba “estrella” para muchos docentes, y cuestionada por otros. Persiste como tal en los DELE y los certificados de las EEOOI de Cataluña, pero no en los de las EEOOI de la Comunidad de Madrid.

La primera observación que podemos formular es que en un proceso de varias horas, que aborda diversos tipos de documentos y propone un importante número de tareas, receptivas y productivas, EL DISEÑADOR Y EL CORRECTOR DE PRUEBAS DEBERÍAN ENCONTRAR INFORMACIÓN SUFICIENTE PARA ESTABLECER UNA MEDICIÓN DE LA COMPETENCIA GRAMATICAL DEL CANDIDATO SIN RECURRIR A UN APARTADO / EXAMEN ESPECÍFICO. En este sentido, ya hemos sugerido la posibilidad de intercalar ítems explícitamente dedicados a comprobar el conocimiento / competencia gramatical. A ellos habría que añadir datos de baremación detallados para aquellos entornos en los que los ítems no se vinculan a lo gramatical pero cuya intervención está implícita (un texto destinado a comprensión escrita con una presencia marcada de, por ejemplo, estructuras condicionales que el candidato es capaz de resolver adecuadamente aunque no afronte ningún ítem concreto, explícito, sobre las oraciones condicionales, ha de constituirse como informador de ese conocimiento / competencia).

En segundo lugar, la observación, también obvia, de que el trabajo gramatical se suele validar mediante textos escritos exclusivamente y éstos ametrallados por huecos. El proceso de lectura escapa a la normalidad del acto de adquirir información. De mantener ese esquema, no parece difícil y sí sensato ofrecer un pequeño resumen del contenido del texto, imposible de leer.



Captura de una de las pruebas de gramática de una de las muestras del DELE correspondiente a noviembre de 2008.

Técnicamente la ejecución es asimismo complicada: texto, opciones o enunciados y finalmente hoja de respuestas: tres planos en los que moverse en los que el primero es fragmentario. También habría que poner solución a eso, porque entre otros efectos indeseados en ese circular entre hojas la gramática aparece casi desprovista de significado; sus relaciones no rebasan el cotexto más inmediato, que es el que por otra parte suele condicionar los ítems (selección de preposiciones, de pronombres y del sistema verbal). De esa formulación no se puede concluir que el candidato haya comprendido el enunciado y el papel que esas formas lingüísticas desempeñan en la construcción del mismo, lo que nos ha de hacer reflexionar cuál es el papel que le es asignado a la gramática.

Y por supuesto, como ya hemos señalado, las técnicas empleadas son terreno abonado –en lo gramatical y en lo léxico– para las ambigüedades y las imprecisiones. Así lo ilustra el siguiente ejemplo.

27. – Cuando fui a recoger el coche me habían puesto una multa.
– ¡Qué **faena**!
a) raro
b) bien
c) pena
28. – ¿Qué te pasa? Pareces preocupado.
– Es que no soy capaz de comprender cómo se maneja este programa. Es **un lío**.
a) muy difícil
b) muy confuso
c) muy lento
29. – ¿Ya has **arreglado** lo de las vacaciones?
– Bueno, no del todo.
a) comparado
b) reservado
c) solucionado

Muestra correspondiente a la edición B1 de mayo de 2009, en la que se aprecia la indeterminación de las propuestas ofrecidas: la 27 no admite ninguna de las opciones sugeridas (la esperable, próxima a la intención de cooperar expresando solidaridad ante la situación adversa sufrida por el emisor, sería algo próximo a *¡qué putada!* o algún impropio dedicado a los agentes); para la 28, son sensatas tanto la a) como la b), cuyas fronteras semánticas son difusas; y, finalmente, son admisibles varias opciones en función del contexto de aparición de los enunciados.

En cualquier caso, es sobresaliente la gran distancia del tratamiento de la gramática (y del léxico) en el aula y en situación de examen. Una suerte de mutación para la que, paradójicamente en términos de aprendizaje, hay que preparar, entrenar específicamente, a los candidatos.

Si abundamos en la decisión de que una prueba de medición nos resulte informativa, abogaremos de nuevo por la posibilidad de elegir una segunda

opción, entre un repertorio más o menos próximo, para determinar el tipo y la calidad de error (e incluso la razón) del error cometido.

EXPRESIÓN ESCRITA

Temida y odiada por muchos estudiantes; la expresión escrita actúa a menudo como validadora -o no- de una competencia comunicativa en un sentido amplio, como si en el fondo los correctores mostraran desconfianza hacia la capacidad de medición obtenida en las pruebas “objetivas”. En el fondo no les falta razón, pues se trata de lengua en uso, de verdadera competencia, rebasando el concepto de conocimiento (aunque se puedan desplegar estrategias –muchas de ellas miméticas- para la redacción). Como sucede con la interacción oral, no se presentan ítems de respuesta (no en general), sino propuestas de expresión, en algunos casos con alguna alternativa al trabajo (dos opciones para desarrollar una, o similares), lo cual parece muy razonable.

El problema fundamental son, claro está, esas propuestas. La mayoría de esas apuestan por la asunción de un *role-play*, que enmarca la situación, los contenidos y el propósito comunicativo. Y ello SUPONE QUE EL PROCESO DE REDACCIÓN SE DEBE ACOMPAÑAR DEL DE IDEACIÓN, QUE NO ES NECESARIAMENTE UNA COMPETENCIA QUE DEBAMOS EVALUAR ni un componente inherente a todos los estudiantes (Alderson, Clamphan y Wall: 1998). Otras instrucciones, en especial en niveles intermedios y avanzados, se orientan a la abstracción, temas abstractos, posiblemente para calibrar la competencia en la expresión de opiniones, los razonamientos y la argumentación, notablemente complejos (Nunan: 1996), ciertamente, también en la L1, dando por válido el axioma de nivel avanzado igual a tema complicado. Tal tipo de propuestas parecen dar por sentado un nivel intelectual, de conocimientos, que el candidato no tiene por qué poseer ni siquiera en su L1. Por fin, las propuestas pueden entrar en colisión con el objeto deseado de medición, ya sea por una formulación inadecuada o por unas instrucciones imprecisas... Así, por ejemplo, en una reciente prueba de certificado elemental (A2) se pedía a los estudiantes que *relataran* el modo en que “internet ha cambiado tu vida”; para el redactor de la prueba la palabra “relatar”, que en efecto aparecía en el enunciado, parecía suficiente para que los candidatos hicieran intervenir en su texto un repertorio pertinente de tiempos de pasado, que era uno de los objetos del nivel y de la prueba; sin embargo, el término entraba en conflicto con “ha cambiado”, de manera que la mayoría de los textos resultantes fueron relatos en presente de indicativo sobre las consecuencias actuales de la internet en sus vidas: la tarea estaba cumplida, pero la prueba como instrumento de medición fracasó, y con ella se descompensó el conjunto del examen (y en algún caso pueden no existir otros elementos medidores). Como es natural, la solución no parece ser la de añadir datos explícitos que conduzcan hacia los exponentes deseados (en rigor, no podríamos calificar ese producto como expresión escrita), sino insistir en el cuidado de las situaciones y de las instrucciones. Con todo, y puesto que nos hallamos ante una prueba de medición, entiendo que es preferible establecer cuáles son los contenidos

esperados en el nivel, los que se persigue medir, que formular propuestas situacionales rocambolescas.

Por nuestra parte, hemos ensayado un par de medios con los que dar lugar a la aparición a los exponentes deseados (por general típicamente gramaticales) en un contexto de “normalidad”. Son las REDACCIONES BREVES ASOCIADAS A OTRAS PRUEBAS que ya hemos mencionado; si el texto o la audición han logrado despertar su interés, los candidatos sienten necesidad de expresar su reacción ante la información recibida (comparar lo explicado con un caso semejante, aconsejar a alguien que ha relatado un problema, resumir o parafrasear un texto...). Estos pequeños textos pueden asimismo ser capacitadores de una tarea final en la que tengan cabida exponentes próximos. Así lo hice con el siguiente ejemplo:

20.

Relee las líneas 26-30. Imagina la situación si Obama decide irse con el hombre de Senegal y los cambios en su vida (25-30 palabras).



Como remate a un fragmento de un libro autobiográfico de Barack Obama, en concreto un viaje a Barcelona con un emigrante africano (el tema sorprendió e interesó a los aprendientes), se les propuso una tarea (B1+) que proponía la utilización del condicional (simple o compuesto). La breve extensión solicitada (limitada visualmente por el tamaño de la imagen, algo que inconscientemente cumplieron) y la emotividad del relato (además de sus conocimientos sobre Obama) dieron lugar a que no redactaran varias frases sino una, con cierta expansión sintáctica. El ítem puede aportar información para la comprensión y la expresión escritas y también para la competencia gramatical (uno de los objetivos perseguidos en casi todas las tareas de expresión escrita).

Además de proponer diversos textos cortos a lo largo de las pruebas, hemos indagado la presentación de un *input* base para el redactado. No es, por supuesto, una técnica nueva (es muy razonable el uso que se hace en algunos de

los exámenes de Cambridge, por ejemplo), pero seguimos explorando sobre ella. Básicamente estos textos (siempre escritos, pero podrían ser documentos orales, algo inferiores al nivel), pueden adoptar tres funciones:

a) objeto de respuesta: el candidato recibe una carta que debe contestar. Se acumulan aquí situaciones y redactados notablemente torpes, ajenos a las máximas de cantidad y calidad, rematados por un entusiasta ¡*Cuéntame!* En el mejor de los casos, al candidato le basta con reescribir el texto original en una operación de muy bajo riesgo. No es necesaria una comprensión completa del texto recibido.

b) El texto como condicionador de los contenidos. Es el caso de la típica carta de reclamación sobre un producto o servicio, que se ilustra con un extracto de las garantías cubiertas y las condiciones (ejemplo que aparecía en una de las muestras del *Certificate of Advanced English*), y que se puede adaptar a muy diversas situaciones.

c) El texto como estímulo de la producción escrita. Texto con interés como para opinar, debatir o continuar... En uno de los modos ensayados con grupos de B1+, se presentó a los candidatos un reportaje periodístico medianamente extenso sobre un fotógrafo de guerra como trabajo de comprensión escrita, con los ítems habituales más algunos de los que se han sugerido hasta ahora. Teníamos la certeza de que el texto le resultaría interesante porque ya se había probado con anterioridad. Tras completar esa parte del examen, los candidatos tuvieron una pausa de unos 20 minutos, en la que en muchos casos hablaron de las partes del examen, incluido el reportaje. Después de la pausa, iniciaron la parte de expresión escrita: se les entregó nuevamente el mismo texto (que conocían bien, incluso había tenido oportunidad de hablar sobre el mismo). La instrucción que recibieron es que debían explicar ese texto a un amigo español, su opinión sobre el mismo y justificar qué hubieran hecho si hubieran estado en la situación del fotógrafo (entre 200 y 250 palabras) y se les informó que el texto del reportaje les sería retirado a los diez minutos exactamente. Eso dio lugar a uno de los objetivos que esperábamos, la confección de borradores sólidos, trabajos, con diversas estrategias (mapas conceptuales, esquemas, anotación de términos de difícil ortografía, minirresúmenes de cada uno de los párrafos, etc.). Más allá de lluvias de ideas (a veces chaparrones que únicamente encharcan el texto), los candidatos consiguieron armar un proceso sólido (y natural: explicar y opinar sobre algo que habían leído, y por consiguiente con cierta capacidad para abstracciones) que se tradujo en buenos o muy buenos productos finales, todos ellos mucho más extensos de las 200 o 250 palabras solicitadas. Es cierto que hubo casos de copia literal del original, algunos inaceptables, pero en otros la copia servía como sostén para generar discurso nuevo, y no tenía por qué ser necesariamente penalizable. Aunque con las limitaciones propias de una situación de examen, trasladamos así modos de operar de aula (la confección de borradores), que entendemos indispensables para la confección de un texto escrito cohesionado y coherente... Hemos seguido explorando esas posibilidades, en nuevas ocasiones con vídeos breves (con un primer visionado sin sonido destinado precisamente a que el objeto de atención fuera lo visual

para activar hipótesis y vocabulario), y los resultados han sido a mi parecer muy razonables, tanto en proceso como en producto. Entre otras, las consignas que se le ofrecían al candidato eran la esperable: que explicara que acababa de ver un vídeo¹¹ (a modo de apertura de discurso), su impresión acerca del vídeo, lo que la persona decía, la caracterización de la misma a partir de lo visto y oído y los motivos por los que le había gustado o no (expresión de opiniones y justificación, comparaciones...), consignas que cubrían los aspectos que queríamos medir.

Permítaseme un pequeño apunte sobre este asunto de la expresión escrita: el escaso equilibrio entre las exigencias de la expresión escrita en un examen y la inversión en el aula, en términos no de corrección lingüística sino de corrección textual... Ciertos estudiantes pueden transferir esas habilidades desde su L1, pero muchos otros carecen de ella. Podemos enseñar lengua, pero nos resulta imposible enseñar a “escribir”, a dotar de coherencia, cohesión, pertinencia y adecuación a los textos de algunos de nuestros estudiantes, y mucho más en la medida en que lo deben hacer en un código que no dominan. Corregimos, y mucho, comentamos, aconsejamos, pero ese trabajo es del todo insuficiente para muchos estudiantes, buenos estudiantes. Tal vez necesitaríamos más tiempo en la clase (la verbalización se considera más eficaz que la corrección “tradicional” –Brender: 1998- para la producción escrita) y tratamientos individualizados, acordes también con trazos asociados a nacionalidades. Sin embargo, los niveles de exigencia son idénticos para todos ellos..., incluso superiores a los niveles de competencia que pueda alcanzar una estudiante que finalice el bachillerato español...

Por otra parte, los criterios de corrección (¿valoración / medición?) son confusos para los candidatos y ambiguos en muchos casos para los correctores. Criterios (Eguiluz y De Vega: 1996 / 2009) como adecuación, estructura o riqueza lingüística (piénsese, por ejemplo, que para un aceptable redactor anglófono los conceptos de complejidad sintáctica o de repeticiones léxicas no poseen el mismo sentido que para un redactor medianamente competente de español peninsular). A la práctica, en muchos casos, esos criterios sustentan la penalización: se incide más en los errores, en lo que el candidato no sabe hacer que en sus logros, a menudo eclipsado por su interlengua escrita.

En mi opinión existen pocas alternativas más para las producciones escritas... Desaparecidas por completo las basadas en la traducción (cuyo papel se ha de reasignar (Cantero y De Arribas: 2004), pues los estudiantes siguen practicándola, ahora “a escondidas”), sólo perviven algunas muestras insólitas, comúnmente alojadas en niveles iniciales, donde se percibe casi exclusivamente un interés en las transformaciones morfológicas... Por lo demás, ninguno de los casos propuestos se formularía por escrito en situación real.

¹¹ El vídeo en cuestión, con un cierto tono lírico es el que se puede ver en http://www.youtube.com/user/nachorpedra#p/u/3/_kUPwQLdGqQ

Queda deliberadamente en el aire la posibilidad de abordar una prueba de expresión escrita con apoyos en diccionarios y materiales de referencia: los recursos con los que un hablante de una L2, incluso con una competencia alta, aborda una tarea de escritura de importancia¹². Estas pruebas existen en algún sistema educativo, pero no siempre suelen estar bien administradas...

EXPRESIÓN / INTERACCIÓN ORAL

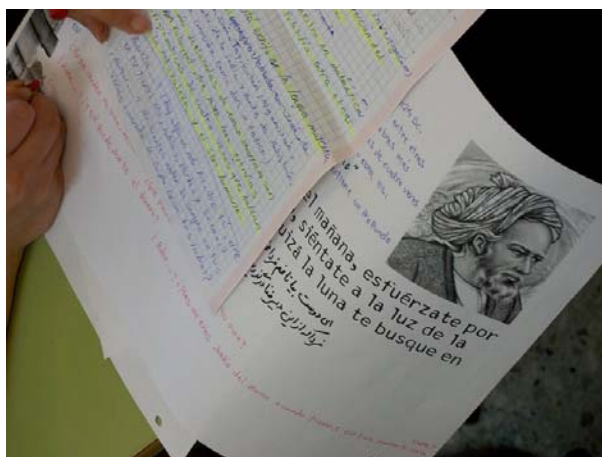
Las opciones de supuestos se repiten en este tipo de exámenes, que adoptan modalidades diversas en relación con los candidatos y la intervención del tribunal. Se combinan organizaciones con uno, dos y hasta tres estudiantes conversando entre sí, con o participación del tribunal examinador. Sobre este aspecto, subrayar ya inicialmente que no resulta justificable esta prueba sin INTERVENCIÓN DE UN HABLANTE NATIVO / COMPETENTE, por muy diversas razones, pero en especial como instrumento capital para disponer de instrumentos de medición vinculados a la comprensión auditiva en un marco de cierta espontaneidad, para amortiguar los efectos de una situación comunicativa anómala: el caso de dos hablantes que comparten L1 conversando en una L2, y también para calibrar fenómenos de interacción auténtica de orden sociopragmático (turnos de palabra, cohesión conversacional, interpretación / producción gestual...) que no son detectables en las interacciones entre iguales, además, claro está, de provocar en alguna medida la aparición de los exponentes esperados en el nivel con aceptable naturalidad.

Las pruebas se resienten de una escasa originalidad técnica. Es inevitable la aparición de una fotografía que el candidato debe describir. Dos problemas asociados a ello: por un lado, que a menudo se espera que la descripción propicie la aparición de un sinfín de exponentes lingüísticos, más allá de lo que suele / puede dar de sí una descripción (no mucho más allá del presente de indicativo, adjetivación y localización espacial). Parece, en efecto, que los redactores de pruebas consideran que el estímulo visual activa necesariamente la producción oral, pero para ello –el segundo problema– la selección de fotografías debiera responder a algo más que plasmaciones espaciales: actitudes, gestos, acciones previsibles (a pesar de ser la fotografía una imagen estática), situaciones insólitas... pueden promover descripciones, y suposiciones y reacciones y opiniones... Determinadas fotografías, simples e inexpresivas, solo son capaces de generar por sí mismas (al margen de las intervenciones del examinador) descripciones estáticas, salvo que las estrategias (léxicas y "argumentales") del candidato estén francamente desarrolladas... Por lo demás esas fotografías suelen llevar aparejados unos temas por lo general manidos: la contaminación, la escuela, la familia, el tráfico, las ciudades, la alimentación..., cuyos resultados pueden ser sólidos (se trata de temas con frecuencia abordados

¹² La trayectoria de producción escrita de un estudiante de una L2 se ha desarrollado en casa, con los apoyos mencionados. Seguramente la situación de examen es de las pocas en las que prescinde de ellos.

en el curso), pero probablemente poco espontáneos (por previsibles), ajenos al discurrir de una conversación normal.

Postulamos un procedimiento que determine los contenidos objeto de medición, que no han de ser necesariamente “argumentales”, y con ese planteamiento establecer la selección de fotografías, entendiendo que el número de exponentes observables es, en principio, limitado (aunque posiblemente redundante).



Para un examen local, interno, propuse una foto como ésta y otras similares, que correspondía a una reciente actividad que realizamos en clase. La imagen tuvo la capacidad de evocar recuerdos (y sensaciones, y tal vez opiniones) e ingredientes para la aparición de los tiempos de pasado (en especial el indefinido), que era el objeto de la medición. La capacidad para rememorar es una de las cualidades esenciales de la fotografía, pero no acaba de ser explorada.



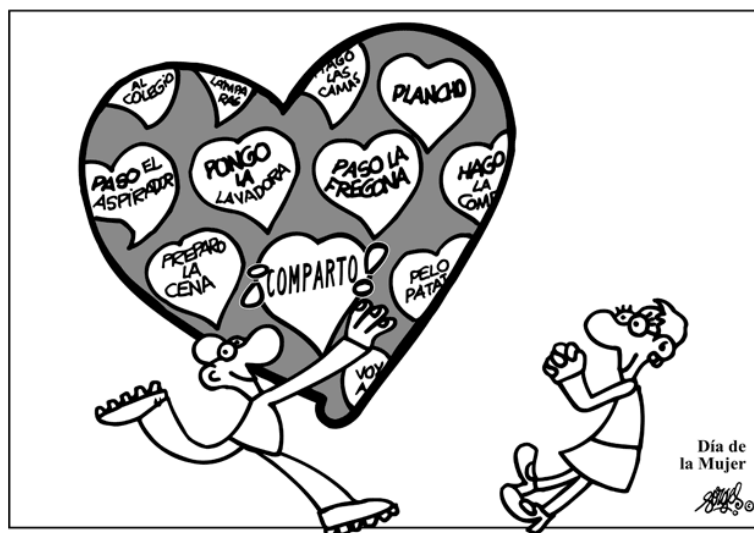
La descripción estática y la expresión de acciones futuras o de suposiciones podrían estar asociadas de forma espontánea a imágenes como éstas.



La opción de usar el pretérito presente de indicativo y el pretérito perfecto compuesto parecen consustanciales a la descripción de un fotografía como la que se reproduce.



Las descripciones más "complejas", que incluyan comparaciones entre dos fotos o con las fotografías de un compañero, a modo de vacío de información, pueden producir un discurso más rico (en *El Ventilador* se encuentra alguna propuesta que parte de este tipo de imagen).



La inclusión de algún tipo de texto, como el que se recoge, puede resultar estimulante en la medida que apuntala el discurso del aprendiente con ciertos exponentes en los que apoyar su intervención. El examinador deberá valorar la capacidad del candidato para retomarlo y construir discurso nuevo. Sin embargo, la inclusión de texto ha sido frecuentemente rechazada por ser considerada como una "ayuda": las interacciones orales se construyen, precisamente, con la cooperación del interlocutor.

Junto con las fotografías, de nuevo los supuestos situacionales, que persiguen la producción de los exponentes lingüísticos que se pretende medir. Los problemas, similares a los ya comentados: la pertinencia esas situaciones para los aprendientes con arreglo a su diversidad cultural (hemos visto algunos ejemplos) y también para la medición de la tarea misma: no bastan unas instrucciones para la producción oral de aquello que se desea medir.

Es posible revertir en parte la concepción de las situaciones. Frente a planteamientos a veces insólitos, situaciones en que el candidato debe simular un rol que le es ajeno, que carece de interés para él o que no es cómodo para una interacción espontánea, tal vez es posible articular otros modos de operar, que de algún modo subrayen la situación de examen, esto es, que sugieran que hay unos exponentes de medición esperados¹³. Y eso supone reconocer que no es posible “medir” todo el contenido del curso. Ni seguramente es necesario.

Recientemente ensayamos algunas fórmulas que, bien administradas, proporcionan resultados correctos. Una de ellas fue la de articular el trabajo con fotografías a un vacío de información (en A2, B1). El desarrollo óptimo es con una pareja (aunque en niveles avanzados se puede organizar con un grupo de tres candidatos). Cada candidato dispone de una fotografía, seleccionada con los criterios mencionados arriba, ligeramente diferente a la de su(s) compañero(s). El objetivo, al describir la imagen, es hallar un número determinado de similitudes y de diferencias entre las fotografías. Inherente al proceso es que los candidatos escuchen al compañero y que procesen la información recibida para así resolver la tarea. Comúnmente la actividad ha dado lugar la petición de aclaraciones, corroboración de lo oído, confirmación de los progresos en la tarea, asunción de turno de palabras, etc.; a negociación de significado en buena medida. Entre los procedimientos analizados, es el que nos ha brindado muestras de lengua más espontáneas y auténticas, siempre en relación con el nivel. Con buena concentración por parte de los examinadores, es posible captar rasgos de estrategias conversacionales, sociopragmáticos y también prosodémicos al ser el segmento más natural de la interacción.

El procedimiento se reproduce en parte con textos breves (a modo de pie de foto), también diferentes pero con un tema próximo. El objetivo es de nuevo natural: transmitir una información leída y compartir esos contenidos próximos para generar un nuevo discurso. Como es obvio, la selección de los textos (y también de las fotografías) es capital para recuperar algunos de los exponentes que se pretende validar. En cualquier caso, y como hemos dicho, el conjunto de los exámenes debe haber procurado oportunidades suficientes para constatar el grado de dominio.

¹³ La prueba / examen de interacción oral, al igual que sucedía con la de expresión escrita, tiene unos objetivos de medición por parte de los examinadores y también un alto componente de producción abierta por parte de los candidatos. Es necesario acercar ésta a aquéllos, aunque sea de modo medianamente explícito. De otro modo, la prueba resultara fallida en su calidad de medidora de una determinada competencia.

Alimentación y salud

Candidato A



Según un estudio realizado en los Estados Unidos, el consumo habitual de comida basura aumenta el riesgo de sufrir obesidad. Los expertos han analizado durante 15 años los hábitos alimenticios de tres mil personas y su repercusión en sus condiciones físicas. Así, las personas que van dos días a la semana a restaurantes de comida rápida han engordado 4,5 kilos más que las personas que lo hacen sólo un día.

Extraído y adaptado de:
<http://www.loquesomos.org>

Candidato B



La obesidad es una enfermedad grave que actualmente se soluciona con una operación que dura dos horas. Este tratamiento es muy agresivo porque el problema es muy importante. Si un adulto no ha aprendido a comer de forma sana y mantener un peso adecuado y estable, no lo hará con dietas que exigen voluntad y sacrificio. Con esta operación tiene la garantía de perder peso y no volver a ganarlo.

Rami Mikler: Una cirugía para perder peso.
<http://www.susmedicos.com/obesidad.html>

1. Observa tu fotografía y pregunta a tu compañero sobre detalles de la suya para encontrar diferencias y parecidos entre las dos.
2. a) Explica lo que dice tu texto y relata alguna experiencia tuya o de alguien que conozcas, relacionada con el tema.
b) Habla con tu compañero sobre el tema presentado por vuestros textos, opina y muestra acuerdo o desacuerdo de forma argumentada, propón soluciones a la temática que estáis discutiendo.
3. Selecciona dos o tres temas y haz una pregunta a tu compañero; responde a sus preguntas sobre estos mismos temas:
 - sus hábitos con la alimentación (dieta, horarios, lugar...),
 - sus gustos,
 - la preparación de algún plato,
 - la preparación de algún plato,
 - un recuerdo concreto relacionado con la comida.

Para terminar, el profesor puede hacerte alguna pregunta relacionada con el tema.

Muestra de la nueva prueba de expresión oral de las EOI de Cataluña (tomada de http://phobos.xtec.net/eoi/index.php?module=detic_portal&ref=es_el_eo en septiembre de 2009). Además del trabajo que se describe, la tarea propone otras actividades, que seguramente dan lugar a un conjunto demasiado ambicioso, que no brinda los objetivos perseguidos.

También fundamental, el papel del examinador para conducir el discurso. Entre las instrucciones presentadas a los candidatos se apuntaban algunos aspectos sobre los que podría ser preguntado. En concreto, en B1+ anunciábamos que a los candidatos se les podía solicitar un resumen de lo que había dicho el compañero (una posible forma de revisar el estilo indirecto), su opinión, etc. Pero la prueba no iba, ni quería ir, más allá de lo que de forma natural podían dar lugar los estímulos ofrecidos.

No podemos concluir este apartado sin recordar, como sucedía con la expresión escrita aunque aquí incrementada por lo efímero de la producción (en los casos en que ésta no es grabada), el carácter etéreo de las pautas de medición, expresadas muchas veces con enunciados vagos (y penalizadores) y con frecuencia desconocidos para los candidatos.

ALGUNA ALTERNATIVA: PARA EL PRESENTE, SIN CIENCIA FICCIÓN

Han aparecido hasta ahora algunas propuestas que se mostraban comprometidas por la propia administración de los exámenes, y hemos aludido en alguna ocasión la configuración de nuevos entornos. Los documentos curriculares actuales insisten en la importancia del aprendizaje para *la acción*, un concepto difícil de preservar en las situaciones de medición, sobre todo si se desea mantener un perfil académico y planificado de las pruebas. Se podrá esgrimir que los exámenes no son una herramienta adecuada para la medición de competencias, y que el futuro ha de prever otras formas de verificación. Probablemente sea cierto, pero no lo abordamos ahora. Nos limitamos a algunos apuntes para el presente más inmediato.

En efecto, la configuración de pruebas en un soporte informático puede ofrecer alternativas plausibles a esos problemas anunciados y/o intuitos, considerando sobre todo que nos hallamos ante unas situaciones académicas en las que se limita o se impide la cooperación entre aprendientes.

Dejando de lado los planteamientos especialmente conductistas que han presidido las acciones con la intervención de ordenadores en la enseñanza de lenguas (sería una tentación demasiado fácil de incurrir en muchos de los errores que hemos apuntado, a pesar de que un buen número de ejercicios podría tener cabida en esas plataformas), es posible articular un modelo de examen significativo con soporte electrónico. Existen diversos programas de autor que pueden albergar muchos de los requerimientos que hemos señalado; también es posible la confección con formato .html o similares¹⁴. No obstante, la manipulación de esos programas o esos lenguajes de programación no siempre son demasiado fáciles para los no profesionales; la inversión en tiempo es alta y la flexibilidad, irregular.

Las posibilidades más al alcance de cualquier docente son las que puede brindar el programa *Acrobat Pro*, en especial su versión 9, que puede partir de un documento previo confeccionado con *Microsoft Word*. En efecto, la versión citada incluye diversas prestaciones una vez creado el documento *Word*. Entre ellas, la inclusión de audio o vídeo (manipulable o no) en cualquier tipo de formato, campos de texto abierto, opciones de selección, etc.¹⁵

Tales planteamientos harían posible, entre muchas otras cosas, la individualización de los trabajos de comprensión auditiva, la inclusión de preguntas orales en pruebas no necesariamente de comprensión auditiva, la ubicación en un mismo plano visual de los diferentes constituyentes del examen, etc. El conjunto de las pruebas podría estar disponible en una red local de acceso

¹⁴ El test de nivel que se recoge en la portada del *Aula Virtual de Español* es un ejemplo de esas posibilidades, en este caso desarrolladas en un entorno Flash:

http://ave.cervantes.es/prueba_nivel/default.htm?Idioma=esp

¹⁵ Parte de las pruebas desarrolladas por Casañ –véase la nota 9- adoptaron este formato.

restringido / limitado, donde el conjunto podría ser depositado, una vez completadas todas ellas. Este tipo de acceso no comportaría problemas con el peso de los archivos (con sus componentes de audio o vídeo), algo que sí incidiría negativamente con accesos mediante conexión a internet.

La administración de la prueba requeriría únicamente de un ordenador personal (tal vez una máquina portátil con conexión inalámbrica) para cada uno de los estudiantes, algo que empieza a ser normal en el panorama académico occidental, unos cascos y el programa (gratuito) *Acrobat Reader*. El proceso automatizado se tornaría amaneuse en lo que respecta a la corrección; a pesar de que no sería estrictamente necesario para los ítems de corrección "objetiva", la creación de una base de datos que retroalimentara la corrección supondría un trabajo ingente y muy profesionalizado: un esfuerzo probablemente excesivo, considerando además que el diseño aquí presentado contempla un número relevante de ítems abiertos. Dejemos lugar para el factor humano.

Con todo, subrayar que no estamos hablando de soluciones, sino de alternativas. Hemos insistido en lo que parecen los aspectos más beneficiosos y flexibles, los problemas hacia una transición parcial hacia test informatizados no son menores... (véase al respecto Dooley, 2008, quien, sin embargo no sobrepasa la simple transferencia de un soporte a otro, sin aproximaciones a lo multimedia, por ejemplo).

ALGUNAS CONCLUSIONES

Parte los contenidos vertidos en este trabajo inciden sobre el mestizaje de técnicas y formulaciones híbridas, con oportunidades de reaparición. Tales planteamientos exigen un diseño minucioso y al mismo tiempo con una capacidad de perspectiva sobre todos los procedimientos de medición. El proceso es laborioso pero no extremadamente complejo y podría ser resumido en una tabla como la siguiente:

CONTENIDOS OBJETO DE MEDICIÓN	INTERACCIÓN ORAL	COMPRENSIÓN ORAL	COMPRENSIÓN ESCRITA	EXPRESIÓN ESCRITA	¿GRAMÁTICA Y LÉXICO?
CONTENIDO ₁ Descriptor(es)	reconocimiento / producción	reconocimiento / producción	reconocimiento / producción	reconocimiento / producción	reconocimiento / producción
_____	objetivo / subjetivo	objetivo / subjetivo	objetivo / subjetivo	objetivo / subjetivo	objetivo / subjetivo
_____	puntuación + / - redundante	puntuación + / - redundante	puntuación + / - redundante	puntuación + / - redundante	puntuación + / - redundante
CONTENIDO ₂ Descriptor(es)					

CONTENIDO _n Descriptor(es)					

TOTALES NUMÉRICOS					

La primera decisión didáctica¹⁶ –y de extraordinaria magnitud- es la determinación de los contenidos objeto de medición, con sus correspondientes descriptores, lo que posiblemente hará innecesario un examen específico de gramática y léxico. En ese detalle deberían tener cabida determinados aspectos socioculturales.

El contraste entre descriptores y respuestas finales permitirá recopilar información relevante para el corrector, y por supuesto para el candidato si, como parece lógico, se le brinda algún tipo de retroalimentación más allá de guarismos (en ocasiones los exámenes se muestran como actividades de información oculta para los estudiantes). El procedimiento de reaparición permitirá asimismo aumentar la fiabilidad y determinar si se trata de una competencia presuntamente consolidada. A simple vista podemos verificar si el contenido deseado está presente en los documentos y en las actividades presentadas y en qué modo: la cantidad de lengua aportada no coincide por supuesto con la cantidad de lengua “preguntada”, pero ello no significa que no sea lengua manipulada en algún modo para la resolución de las pruebas. Podemos igualmente organizar de inmediato el tipo de ítem para cada caso y así determinar los modos en que se mide (productivos o no). Finalmente la observación del descriptor, su eventual grado de reaparición y el modo de medición deberían establecer la puntuación asignada: no hay razón, en principio, para cuantificar un mismo descriptor más de una vez (a lo sumo determinar su grado de consolidación), ni tampoco, por ejemplo, mantener mediciones numéricas iguales para contenidos diversos (revisión / presentación) o de más difícil adquisición (aquí podrían entrar en juego variables individuales contrastivas o eventuales ítems opcionales), e incluso las mencionadas “segundas opciones” para ítems de selección múltiple en los casos en que se haya determinado un repertorio significativo y discriminador adecuado. El objetivo final de una medición más fiable y más informativa es el que preside el conjunto.

Obviamente la dificultad estriba en la selección de textos (escritos y orales) para proceder a esas mediciones, pero la formulación es, inicialmente, la adecuada: establecer qué es lo que debemos buscar, no esperar que el azar provea un batiburrillo de muestras de lengua.

Tal vez se considere que muchas de las ideas aquí formuladas sean ingenuas. Seguramente sí. Suponen ciertamente la modificación de muchos hábitos anclados en el tiempo, que exigen además decisiones institucionales significativas. Ingenuas. Seguramente sí, pero lo que sería calamitoso es que carecieran de sensatez. Lo que corresponde ahora es procurar contextos para la investigación y el análisis de resultados de algunas de las propuestas, verificar su fiabilidad, su afectividad y su capacidad de información.

¹⁶ Una formulación como ésta –o similar, pues sin duda es francamente mejorable- es la que debería presidir los criterios de corrección de las pruebas, al margen de apuntes más o menos subjetivos.

REFERENCIAS BIBLIOGRÁFICAS

Todos los enlaces electrónicos se mostraban activos en septiembre de 2009

ALDERSON, J. Charles; CLAPHAM, Caroline; y WALL, Dianne (2001): *Exámenes de idiomas. Elaboración y evaluación*. Madrid, Cambridge University Press.

BORDÓN, Teresa (2004): "Panorama histórico del desarrollo de algunas de las cuestiones fundamentales en la evaluación de segundas lenguas", en *Carabela*, núm. 55 (págs. 5-29). Reedición en *marcoELE, Revista de Didáctica ELE*, núm. 7 (2008), julio - diciembre
<http://marcoele.com/monograficos/evaluacion>

BRENDER, Alan (1998): "Conferencing: An Interactive Way to Teach Writing", en *The Language Teacher Online*; julio
<http://jalt-publications.org/tlt/archive/main.php>

CANTERO, Francisco José y DE ARRIBA, Clara [2004]: "Actividades de mediación lingüística para la clase de ELE", en *redELE, revista electrónica de didáctica ELE*, núm. 2, octubre
http://www.educacion.es/redele/revista2/cantero_arriba.shtml

CÁRDENAS MARTÍNEZ, Abel (2007): "La fiabilidad de los DELE y los candidatos japoneses: ¿un motivo de preocupación?", en *Cuadernos Canela*, marzo, vol. XVIII
<http://www.canela.org.es/cuadernoscanela/canelapdf/cc18cardenas63-76.pdf>

CADIerno, Teresa (2010): "El aprendizaje y la enseñanza de la gramática en el español como segunda lengua", en *marcoELE, revista de didáctica ELE*, núm. 10, enero-junio
<http://marcoele.com/el-aprendizaje-de-la-gramatica>

CASAÑ, JUAN CARLOS (2009): "Didáctica de las grabaciones audiovisuales para desarrollar la comprensión oral en el aula de lenguas extranjeras", *Suplementos marcoELE*, núm 9, julio-diciembre
<http://marcoele.com/suplementos/didactica-de-las-grabaciones-audiovisuales>

CONSEJO DE EUROPA (2002): *Marco Común Europeo de referencia para las lenguas: aprendizaje, enseñanza y evaluación*. Madrid, Instituto Cervantes, Ministerio de Educación, Cultura y Deporte y Anaya.

DOOEY, Patricia (2008): *Language testing and technology: problems of transition to a new era*, *ReCALL* 20(1): 21-34

ESCANDELL VIDAL, M. Victoria (1996): "Los fenómenos de interferencia

pragmática", en *Didáctica del Español como Lengua Extranjera*, Colección Expolingua, Madrid.

Reedición en *marcoELE, Revista de Didáctica ELE*, núm. 9 (2009), julio-diciembre
http://marcoele.com/monograficos/expolingua_1996

EGUILUZ PACHECO, Juan Y DE VEGA SANTOS, Clara M. (1996): "Criterios para la evaluación de la producción escrita", en *Didáctica del Español como Lengua Extranjera*, Colección Expolingua, Madrid.

Reedición en *marcoELE, Revista de Didáctica ELE*, núm. 9 (2009), julio-diciembre
http://marcoele.com/monograficos/expolingua_1996

INSTITUTO CERVANTES (1994): *La enseñanza del español como lengua extranjera. Plan Curricular del Instituto Cervantes*. Alcalá de Henares.

INSTITUTO CERVANTES (2006): *Plan curricular del Instituto Cervantes. Niveles de referencia para el español*. Madrid, Biblioteca Nueva. 3 volúmenes.

KUMARAVADIVELU, B. (2003): *Beyond Methods. Macrostrategies for Language Teaching*. Yale University Press.

KUMARAVADIVELU, B. (2008): *Cultural Globalization and Language Education*. Yale University Press.

LOZANO GRACIA y RUIZ CAMPILLO, J. Plácido (1996): "Criterios para el diseño y la evaluación de materiales comunicativos", en *Didáctica del Español como Lengua Extranjera*, Colección Expolingua, Madrid.

Reedición en *marcoELE, Revista de Didáctica ELE*, núm. 9 (2009), julio-diciembre
http://marcoele.com/monograficos/expolingua_1996

MARTÍN PERIS, Ernesto (1991): "La didáctica de la comprensión auditiva", *Cable*, núm. 8 (noviembre de 1991, págs. 16-26). Reedición en *marcoELE, Revista de Didáctica ELE* (2007), núm. 5, julio – diciembre
<http://marcoele.com/la-didactica-de-la-comprension-auditiva>

MARTÍNEZ, Roser (1997): *Conectando texto. Guía para el uso efectivo de elementos conectores en castellano*. Barcelona, Octaedro.

MAYER, Richard (1997): "Multimedia Learning: Are We Asking the Right Questions?" en *Educational Psychologist*, 32(1) 1-19, primavera; págs. 55-71.

MAYER, Richard (2002): "Cognitive Theory and the Design of Multimedia Instruction: An Example of Two-Way Street Between Cognition and Instruction", en *New Directions for Teaching and Learning*, núm. 89, primavera; págs. 55-71.

NUNAN, David (1996): *El diseño de tareas para la clase comunicativa*. Cambridge, Cambridge University Press.

ORTEGA OLIVARES, Jenaro (1995): "Evaluación en el aula de español LE: algunas consideraciones generales", en *VI Congreso Internacional de ASELE Tendencias Actuales en la Enseñanza del Español como Lengua Extranjera II*. León.
http://cvc.cervantes.es/ensenanza/biblioteca_ele/asele/pdf/06/06_0018.pdf

PASTOR, Carmen (2009): "La evaluación de la comprensión oral en el aula de ELE", en *marcoELE, Revista de Didáctica ELE*, núm. 9, julio-diciembre
<http://marcoele.com/numeros/numero-9>

YAGÜE, Agustín (2008): "Clics, tics y didáctica", en *Mosaico*, Consejería de Educación en Bélgica, Países Bajos y Luxemburgo, núm. 22, diciembre.

WAGNER, Elvis (2007): "Are they watching? Test-taker viewing behavior during an L2 video listening test", en *Language Learning & Technology*, vol. 11, núm. 1, febrero; págs. 67-86
<http://llt.msu.edu/vol11/num1/wagner>