

# **Probando, probando... analizando la utilidad de un examen de ubicación para hablantes de español como lengua de herencia**

**Patricia MacGregor-Mendoza**  
New Mexico State University

## **1. Introducción**

El desarrollo de una prueba ha de realizarse en un proceso consciente e iterativo. Uno ha de tener consciencia de las dimensiones que desea medir, las capacidades que son indicadores de esas dimensiones, el propósito por el cual se creó la prueba, verificar la validez de cada ítem respecto a las habilidades examinadas, evaluar y calibrar cada uno en cuanto a calidad, cantidad y dificultad, establecer los tanteos globales que representan los umbrales de capacidad mínima de cada nivel, y conducir una revisión de todo regularmente para verificar que la prueba sigue cumpliendo su función anticipada (Bachman 1990; Rudner 1994; Bachman & Palmer 1996, Hughes 2003, Weir 2005, Piontek 2008).

Sin embargo, en algunos casos las pruebas no pasan por un escrutinio profundo antes de administrarlas. Es decir, se confía en las pruebas y como consiguiente, en los datos que arrojan, sin cuestionar las habilidades que se examinan ni verificar los méritos de los resultados. Confiamos en las pruebas más por motivos de respeto por los autores que las redactaron y/o el hábito de aplicarlas, que por una comprensión de lo que se está examinando y la utilidad de los datos producidos. Este último escenario, en efecto, ha sido el de la prueba de ubicación que, hasta la fecha, sigue en vigor en la Universidad Estatal de Nuevo México.

La prueba original fue redactada por un lingüista responsable y admirado de nuestra universidad vecina, la Universidad de Tejas en El Paso. Contiene 100 preguntas de elección múltiple que en su mayoría contienen cuatro reactivos, aunque hay algunas preguntas que tienen solamente dos y otras que contienen hasta cinco. Se ha de notar que el desarrollo del instrumento, conocido como el “Examen de Colocación en Español” en una versión temprana, data desde el 1979 (Teschner 1983). Representa un esfuerzo pionero y laborioso por parte del autor de intentar tomar en cuenta las diversas habilidades de la población estudiantil a la vez de identificar las necesidades del departamento de ubicar a dichos estudiantes en un programa de estudio. El autor describe estos esfuerzos en términos generales, la creación de la prueba con el fin de separar dos grupos de estudiantes—los nativos de los no nativos—y de tener alguna idea de la capacidad de los estudiantes dentro de cada grupo para así poderlos colocar adecuadamente en diferentes niveles de estudio de acuerdo con sus habilidades (Teschner 1983, 1990). Sin embargo, no nos ofrece la justificación de la selección de temas en que se basaron las preguntas, ni nos presenta una historia del desarrollo de los ítems (los que se rechazaron, los que se modificaron, etc.) ni un análisis detallado de cada uno de los ítems que forman parte de la prueba actual.

A pesar de la ausencia de información sobre el desarrollo de la prueba, la Universidad Estatal de Nuevo México (NMSU) la adoptó como prueba de ubicación en el 1993, y la ha empleado con los mismos fines generales. Por los primeros 10 años se administraba en forma impresa supervisado por profesores de lengua y los mismos calificaban los resultados a mano o por Scantron y los distribuían a los profesores involucrados en los programas de español básico. En 2004 se cambió la administración de la prueba a un formato electrónico realizado a través del internet con acceso abierto supervisado por un especialista de tecnología. Mientras el cambio a un formato electrónico facilitó la administración de

la prueba y la calificación y distribución de los resultados a los examinados, ha provocado situaciones que han afectado la validez de los ítems, la fiabilidad de los resultados, y el provecho de los estudiantes que la presentan. Los problemas asociados con los cambios no salieron a la luz hasta años después:

1. Se redujo la seguridad del examen y por ende, la fiabilidad de los resultados;
2. Se desniveló el número y dificultad de ítems presentados a hablantes no nativos y hablantes de español como lengua de herencia;
3. Se eliminaron 10 ítems que se cambiaron 10 nuevos sin pilotarlos;
4. Se cambió el orden de presentación de los ítems de uno fijo en grado de la dificultad del contenido a un orden aleatorio;
5. Se introdujo una manipulación arbitraria del puntaje para ciertos ítems, y de los umbrales para entrar a los diferentes niveles de clases.

Además, el cambio de administración de la prueba coincidió con una serie de cambios infraestructurales que, en su conjunto, han operado para reducir el número de estudiantes que entran a los cursos de español como lengua de herencia. La revelación de las complicaciones asociadas con la administración electrónica de la prueba, la falta de detalles sobre el desarrollo de la prueba original, la falta de confirmación de la utilidad de los ítems para nuestros fines curriculares, y el descenso en el número de estudiantes que se matriculaban en los cursos de español para hablantes de español como lengua de herencia motivaron este análisis inicial de la prueba.

## **2. Los hablantes de español como lengua de herencia en el sur de Nuevo México**

Mientras la población estudiantil hispanoparlante de UTEP, para la cual se creó la prueba original, comparte varios lazos lingüísticos y culturales con la población del sur de Nuevo México, hay unas diferencias sutiles pero importantes. Primero, hay un contacto sostenido y regular entre los residentes de El Paso y los de su metrópoli hermana al lado mexicano, Ciudad Juárez. Los más de 30.5 millones de visitantes que cruzan anualmente desde Ciudad Juárez a El Paso (U.S. Department of Transportation 2009), ayudan a estimular no solamente la economía, sino también a mantener un vínculo lingüístico activo y dinámico con México.

Semejantemente, la proximidad de UTEP a la frontera nacional mexicana, a una distancia de menos de dos millas, invita a estudiantes de México a formar parte de su alumnado. UTEP cuenta con 8.92% o 1797 de sus 20,154 estudiantes que son actualmente residentes de México (The University of Texas at El Paso 2008). Este dato implica que estos estudiantes recibieron toda o una parte de su formación pre-universitaria en su país de origen donde habrán forjado sus conocimientos de la norma culta lingüística oral y escrita del español como idioma mayoritario.

En el sur de Nuevo México, los hablantes de español como lengua de herencia (HLH) forman parte de una comunidad bilingüe de diversos orígenes y habilidades lingüísticas. Aunque hay algunas excepciones, típicamente los HLH han aprendido español principalmente de forma oral en un ambiente casero sin haber recibido estudios formales en o sobre el idioma. Sus habilidades se han alimentado por conversaciones familiares, celebraciones culturales, vueltas ocasionales a Ciudad Juárez para ir de compras o a visitar parientes, y las telenovelas y los maratones de películas de Cantinflas, Pedro Infante, y Vicente Fernández que se pasan en la televisión los domingos. Aunque NMSU no distingue entre los estudiantes que vienen de México y de otros países, solo 1,544 (8.35%) de los 18,497 estudiantes provienen de diversos países fuera de los EEUU (New Mexico State University 2010).

Los HLH que mantienen mayor fluidez en su habla en español generalmente provienen de las colonias, pequeñas comunidades rurales que brotan al lado de las carreteras y que carecen de una infraestructura municipal formal. Las colonias están pobladas en gran parte por latinos de diversas generaciones de inmigración, los cuales disponen de escasos recursos económicos y educacionales (Norman, Parcher & Lam 2004). La densidad étnica y las redes sociales de estas comunidades, junto con las posibilidades limitadas de movilidad económica y física de la comunidad, promueven la manutención del español (Pease-Alvarez 1993; Bills et al. 1995). En comparación, los HLH que vienen del área urbana de Las Cruces, como grupo, tienden a exhibir menos facilidad en su habla en

español y mayor uso de inglés ya por experimentar mayor contacto con otras personas que son monolingües en inglés que sus compañeros en áreas rurales. Para todos los HLH, su exposición a la lectura y escritura y al estudio formal del español varía, pero por lo general ha sido limitada o a una escala mucho más menor que su exposición al idioma en forma oral (MacGregor-Mendoza 1998, 2000, 2005).

### 3. Análisis

Para analizar una prueba hay que determinar la utilidad del instrumento de cumplir con su función prevista con la población destinataria. En el caso de una prueba de ubicación, ha de poder identificar y apartar a los estudiantes que son HLH o nativos de los que no lo son, y discernir niveles de capacidad entre ellos.

Para este análisis, examinamos aquí las respuestas de 1610 personas identificadas por la prueba como hablantes de español como lengua de herencia de más de 4700 personas en total que presentaron la prueba durante el año 2007. Este análisis cuenta principalmente con una revisión de los ítems de la prueba en base a principios de la validez. Se sustenta con las estadísticas básicas de análisis de ítems, es decir, los patrones de selección de los diferentes reactivos para las preguntas, y la dificultad del ítem, calculada como la proporción del total de los examinados que acertó la respuesta del ítem.

La validez de una prueba se ciñe a la evidencia proporcionada de diferentes fuentes que, en su conjunto, confirma la legitimidad del instrumento para los fines provistos. Es importante reconocer que la validez no representa una cualidad absoluta que se puede comprobar, sino una colección de condiciones óptimas que se tratan de maximizar. Según la versión más reciente de los *Standards for Educational and Psychological Testing* (American Educational Research Association [AERA], American Psychological Association [APA], National Council on Measurement in Education [NCME], 1999), los pilares de la validez que representan las fuentes de evidencia son: 1) el contenido de las preguntas; 2) las estrategias y los procesos cognoscitivos involucrados en formular las respuestas; 3) la estructura interna de la prueba; y 4) la relación de los resultados a otras variables. Como reflejo de estas consideraciones, hemos adoptado las siguientes preguntas para guiar nuestro análisis de la prueba de ubicación:

1. ¿Reflejan las preguntas las destrezas/los conocimientos necesarios para determinar las habilidad(es) lingüística(s) deseadas?
2. ¿Activan los examinados las estrategias y los conocimientos lingüísticos esperados para responder a las preguntas?
3. ¿Reflejan las preguntas una progresión de habilidad para poder identificar a estudiantes de diferentes capacidades?
4. ¿Sirven los resultados como indicación de las habilidades requeridas para entrar al nivel de estudio a que les mandamos?

Los autores de los *Standards* (AERA, APA, NCME, 1999) señalan una consideración adicional, no para establecer la validez del instrumento, sino para contextualizarlo y dar hincapié a la ética de su aplicación. Piden que tomemos en cuenta las consecuencias sociales de la prueba en su administración y/o en la revelación o uso de los resultados. Es decir, se ha de preguntar si hay efectos secundarios negativos a cambio de administrar o presentar la prueba. También exploramos esta consideración en este trabajo.

### 4. Resultados

Para responder a la primera pregunta miramos el contenido del examen en cuanto a sus dos propósitos: 1) distinguir entre los examinados no nativos y los que sí lo son y 2) identificar diferentes niveles de habilidad. Primero, para separar los dos grupos de hablantes identificados por sus observaciones, nativos y no-nativos, el autor de la prueba incluyó 10 “indicadores de hablantes nativos” (Teschner 1990). Estos indicadores se escribieron con una población mexicana en mente ya que emplean un alto número de coloquialismos típicamente mexicanos. Para determinar los

conocimientos lingüísticos extracurriculares de los examinados se usa la comprensión de expresiones como *recio* para rápido, *greñas* para cabello, *se largó* para se fue, *a poco* para una expresión de duda, y *no se me pegan* para indicar una falta de memoria para datos. Este tipo de pregunta nos permite hacer una división algo bruta entre los que se han expuesto al español exclusivamente por medio de un salón de clase y los que no. Sin embargo, no nos permite una precisión en la distinción entre personas que se han criado en un ambiente de habla hispana y poseen una competencia lingüística más establecida, y los que por una exposición breve pero intensiva a hablantes nativos de español o en el trabajo o por parte de viajes al extranjero, hayan asimilado expresiones selectas de sus interlocutores. Además, se nos presenta la dificultad de determinar cuántas (y/o cuáles) de las 10 preguntas son suficientes para indicar el origen de los conocimientos del examinado y descontar un resultado aleatorio.

Con el cambio de la administración de la prueba se presentaron otros problemas relacionados con los indicadores: el cambio de su lugar en la prueba y la modificación de su uso. En el formato original, estos indicadores se encontraban en medio de la prueba de 100 preguntas que se administraban a todos los estudiantes. Su colocación al interior servía para minimizar la ansiedad que podría desmotivar a examinados que poseen un bajo nivel de conocimientos. En el formato electrónico actual, esos 10 indicadores ahora figuran entre las primeras preguntas de la prueba.

En cuanto a su uso, en la administración anterior, se tomaban estos indicadores en cuenta con los resultados de las otras secciones. Además, en casos dudosos, donde el examinado había acertado entre 4 y 7 de los indicadores, acudía a una entrevista personal con un maestro de idiomas para sondear mejor sus habilidades y ofrecerle una colocación curricular más precisa. En la administración actual, estos indicadores se usan para canalizar a los examinados a una segunda parte de la prueba destinada o para hablantes no-nativos o para hablantes de español como lengua de herencia. Como consecuencia, la decisión, realizada automáticamente, solo accede a los casos extremos. Se ha perdido entonces la capacidad de detectar el rango de habilidades lingüísticas que hay entre la población de HLH y atenderlos con mayor fidelidad y particularidad.

Otro elemento problemático viene a raíz del (mal)uso de los indicadores. Anteriormente todos los que presentaban la prueba recibían las mismas 100 preguntas. En cambio, en el formato electrónico actual que emplea los indicadores como filtro, la segunda parte de la prueba que le toca a los diferentes grupos varía en cuanto a su número de preguntas y en cuanto a su contenido. La prueba para los no nativos consiste en tan solo 50 preguntas adicionales y cubre solo información básica. Contrastivamente, los estudiantes a quienes la computadora clasifica como HLH reciben un total de 90 preguntas más; reciben las mismas 50 preguntas de información básica que los no-nativo hablantes, más otras 40 preguntas de contenido más avanzado. Para todos, las preguntas se presentan de manera aleatoria en el formato electrónico actual. En su formato original las preguntas se presentaban en el orden de las siguientes categorías:

- 1-40 gramática básica (concordancia entre sujetos y verbos y entre adjetivos y nombres, ser y estar, gustar, preposiciones, mandatos)
- 41-50 vocabulario básico (varios nombres y verbos)
- 51-70 vocabulario intermediario y avanzado (varios nombres y verbos)
- 71-80 ortografía (contrastes de **s** y **z**, **g** y **j**, **b** y **v**, etc.)
- 81-90 teoría gramatical formal (identificación de elementos sintácticos, tiempos verbales, etc.)

Para ver si reflejan las destrezas o conocimientos que deseamos examinar, miramos el contenido de las 90 preguntas que forman la segunda parte de la prueba para los HLH. Allí, también hallamos una serie de sospechas. En primer lugar, tenemos que cuestionar la inclusión de elementos básicos como contrastes entre *ser* y *estar* y el uso del verbo *gustar*, que normalmente no reflejan áreas problemáticas para una población de hablantes de español de herencia quienes han adquirido el español en un ambiente familiar en lugar de un entorno escolar. Además, hay que disputar la función de las categorías de dudas de ortografía y teoría gramatical formal que por su lado, representan elementos que, en gran parte, son ajenos a las experiencias de adquisición de esta población de estudiantes.

En cuanto a lo que queda, los elementos básicos, intermediarios y avanzados de vocabulario, el autor de la prueba no nos indica los pasos que tomó para seleccionar los vocablos que incluye en el

instrumento. Tampoco nos informa sobre los criterios que aplicó para determinar la clasificación de su dificultad, ni qué habilidades se revelan por saber una palabra u otra, ni cuántas palabras son necesarias para demostrar una posible competencia léxica, etc. Desafortunadamente, las respuestas a estas inquietudes no son obvias. Así que, a primera vista, varios elementos del contenido no parecen manifestar el beneficio deseado de medir adecuadamente las habilidades lingüísticas de la población de hablantes de español como lengua de herencia y levantan dudas sobre su eficiencia en un examen de ubicación.

La segunda pregunta que propusimos nos pide determinar si los examinados activan las estrategias y conocimientos deseados al presentar la prueba. Para responder a esta pregunta hay que examinar primero la dificultad del ítem para una idea general de la utilidad de distinguir a los examinados por grupos de habilidad. Luego, se explora en más detalle la construcción de las preguntas y los patrones de respuestas que exhiben los examinados al responder a ellas.

La dificultad del ítem representa la proporción de respuestas correctas de todas las respuestas dadas. Una proporción alta, digamos de .85 o más, indica que la pregunta es, a lo mejor, demasiado fácil, ya que la respuesta es transparente para al menos 85% de los examinados sin importar su nivel de habilidad. En ese caso, por extremadamente fáciles, estas preguntas son de poco provecho y por lo general, han de eliminarse ya que no nos ayudan a distinguir entre diferentes niveles de habilidad (Kehoe 1995a, 1995b).

Por el otro lado, las preguntas con una proporción baja, de .40 o menos, han de examinarse para determinar si la dificultad del ítem se debe a la dificultad del contenido (algo favorable en cantidades moderadas ya que ayuda a diferenciar entre grupos de diversas habilidades) o a una deficiencia o ambigüedad en la raíz de la pregunta o en las opciones que tenemos en las respuestas. La proporción óptima para las preguntas varía según el número de opciones que ofrecemos a los examinados. Se calcula dividiendo un resultado perfecto (1.00) por el número de reactivos (en el caso de 4 reactivos, sería .25). Esta cifra representa probabilidad de acertar la respuesta al azar. Se agrega esta proporción (.25) a la mitad de la proporción que quedó (.75/2=.375) de manera que la dificultad óptima para ítems con 4 reactivos es .63; si contiene 5 reactivos es .60, si contienen solamente 2 reactivos, es .75 (Kehoe 1995b). Sin embargo, hay que recordar que la dificultad del ítem es solamente un indicador, no representa una validación de la calidad o aceptabilidad del ítem.

A primera vista, la Tabla 1 revela que al menos 14 de las preguntas del gramática básica y 9 de vocabulario básico son elegibles para eliminarse por demasiado fáciles ya que más del 85% de los examinados las pueden acertar con regularidad. Igualmente hay varias preguntas en las últimas dos columnas, representando un acierto bajo, que merecen atención para determinar a qué exactamente se debe la dificultad de la pregunta. Es decir, hay que tratar de determinar la causa de la falta de acierto para ver si se debe a una mayor dificultad de contenido o si representa un artefacto de la construcción de la pregunta.

Categoría de la pregunta	Dificultad de Ítem	.85-.99	.75-.84	.60-.74	.45-.59	< .44
	<b>1-40 gramática básica</b>		14	8	8	5
<b>41-50 vocabulario básico</b>		9	1			
<b>51-70 vocabulario intermedio y avanzado</b>			2	5	6	7
<b>71-80 ortografía</b>			4	1	3	2
<b>81-90 teoría gramatical formal</b>			1	2		7

**Tabla 1.** Resumen de la Dificultad de Ítem. Frecuencia de preguntas de grado de dificultad según su categoría.

Las preguntas cuyo grado de dificultad cae en las columnas centrales no se salvan de un examen crítico, es solo que son de menor prioridad inmediata. De todas maneras en estas preguntas como en

todas necesitamos confirmar que las preguntas están bien escritas. Es decir, hay que afirmar que reflejan la dimensión que examinamos, y que los examinados activan los conocimientos lingüísticos que esperamos y que estén escritos siguiendo ciertas normas de la formación de preguntas. Entre sus recomendaciones, Kehoe (1995a), Haladyna, Downing, & Rodríguez (2002), Hughes (2003) y Piontek (2008) ofrecen que los ítems han de:

1. Relacionarse claramente con el concepto que se examina;
2. Examinar un solo concepto a la vez;
3. Poseer una raíz adecuadamente clara y libre de ambigüedades;
4. Poseer una sola respuesta correcta que se revela claramente para los que dominan el concepto;
5. Poseer distractores que son creíbles y atractivos para los que no dominan el concepto todavía.

La información en la Tabla 2 presenta preguntas representativas de las diferentes categorías y la calculación de su dificultad de ítem. Además, para cada ítem se documentan las opciones presentadas a los examinados y se señala la respuesta esperada con un asterisco. Se nota también el número de examinados que seleccionó cada opción y/o cuántos omitieron una respuesta. Las negritas marcan las opciones más frecuentemente seleccionadas.

Un examen de estos datos revela varias deficiencias. La pregunta 12, como se ve, representa una pregunta extremadamente fácil, acertada por el 1455 de los 1610 examinados, representando una dificultad de .90. En el otro extremo, vemos varias preguntas cuya dificultad no proviene del contenido sino por una confusión o ambigüedad en las respuestas representan un conflicto de estrategias para los HLH.

Por ejemplo, mientras en la pregunta 11 la respuesta *nos conocen* representa la transformación sintáctica esperada, compite directamente con la respuesta de *nos conocemos* que cumple la misma función pragmática y representa una expresión más natural en el contexto auténtico experimentado por los HLH. Por lo tanto, no nos sorprende ver que 505 examinados hayan seleccionado esta alternativa.

Vemos una contradicción semejante en la pregunta 31. Aquí de nuevo, en base a las reglas sintácticas, se espera la selección de un verbo en el subjuntivo. Sin embargo, la única opción ofrecida en esa conjugación contradice la lógica de los examinados y compite con su experiencia lingüística vivida. En este caso, los examinados, aplicando los conocimientos de su adquisición natural, optan mejor por la alternativa que refleja algo práctico en lugar de sintáctico.

Vemos otros problemas en la construcción de preguntas como números 52 y 64. En ambos casos, la formación de la raíz es confusa de forma intencional. Introducen elementos superfluos pero semánticamente cargados en posición temática (i.e. *el borracho*, *el panadero*) que no tienen función en la oración más que distraer al examinado de la información clave que ocupa para acertar la respuesta. Los patrones de respuestas en la Tabla 2 que se observan para esas dos preguntas revelan que efectivamente, esos elementos desvían a un gran número de los examinados, llevándolos a seleccionar elementos más frecuentemente asociados con un borracho (un trago, la historia de sus amores) o un panadero (un pastel). Por lo tanto, la selección de las diferentes opciones no se puede atribuir claramente a la calidad de los distractores ya que contienen elementos rivales en la raíz.

Por el otro lado, preguntas como la 65, emplean contextos y vocabulario que por lo general no formarán parte de las experiencias vividas y habladas de los HLH. Entre la población de HLH, el término *ferrocarril* se conoce mejor por el sinónimo más común de *tren*. Además, la ausencia completa de trenes pasajeros en el sur de Nuevo México (aunque hay servicio limitado en El Paso), nos obliga a cuestionar lo que la asociación de *ferrocarril* con *andén* puede revelar sobre los conocimientos de esta población de estudiantes para su ubicación eficaz.

Contrastivamente, hay preguntas que parecen penalizar a los estudiantes por sus experiencias vividas y por su forma de adquisición natural. Los ítems 73 y 77 son representativos de este tipo de pregunta. En la pregunta 73, vemos que en lugar de sondear adecuadamente las habilidades sintácticas o semánticas de los HLH, se les pide reconocer la norma ortográfica de una palabra que quizás no hayan visto ni frecuentemente ni consistentemente deletreada. El patrón de respuestas revela que su éxito es aún menos de lo que se podría sacar por el azar simple.

De forma semejante, los conocimientos lingüísticos de los HLH se sancionan en otros ítems. En el 77 por ejemplo, se espera que los examinados escojan entre *nadie* y *nadien*. Dentro de las

comunidades donde los HLH viven y conviven con otros en el sur de Nuevo México, estas formas existen en variación libre. Por lo tanto, no tendrán necesariamente las bases gramaticales formales de una norma culta para poder seleccionar la forma gramatical esperada en una prueba. El patrón de las respuestas confirma que el ítem da raíz a cierta confusión.

Donde se ve la mayor distancia entre los conocimientos lingüísticos de los HLH y lo que se examina en la prueba es en los ítems de la categoría de teoría gramatical formal. El ítem 84, representa este tipo de pregunta donde se les pide a los HLH identificar elementos lingüísticos (generalmente verbos) en base a las formas de sus conjugaciones u otras clasificaciones gramaticales. Como el patrón de respuestas revela, tanto en la Tabla 1 como en la Tabla 2, estos elementos son de bajo éxito para la mayoría de los HLH examinados.

Ítem	Reactivo A	Reactivo B	Reactivo C	Reactivo D	Reactivo E	omitido
12. Tenemos ___ en el invierno. Dificultad: .90	*mucho frío N= 1455	muy frío N=48	muchos fríos N=48	muy fríos N=18		N=41
11. ¿Ellos los conocen a ustedes? Dificultad: .38	nos conocen N=616	los conocen N=239	nos conocemos N=505	los conocemos N=195		N=55
31. ¿Hay alguna persona aquí que ___? Dificultad: .32	puede ayudarnos N=675	llegó antes de la una N=204	*sepa italiano N=526	ha visto el accidente N=147		N=58
52. El borracho me guiñó Dificultad: .43	*un ojo N=703	otro trago N=316	una moneda N=152	la historia de sus amores N=297	las fábricas N=31	N=111
64. El panadero traía ___ en el ojal. Dificultad: .20	un negocio N=66	una patraña N=47	una pezuña N=71	un pastel N=1003	*un clavel N=325	N=98
65. Al llegar a la estación de ferrocarril, fueron directamente al ___. Dificultad: .27	umbral N=232	*andén N=438	escombros N=315	escaparate N=362	yelmo N=105	N=158
73. Mi amigo no ___ interrumpir nuestra conversación Dificultad: .48	*quiso N=772	quizo N=786				N=52
77. Ella no es amiga de ___. Dificultad: .58	*nadie N=943	nadien N=623				N=44
84. Un ejemplo de un verbo que aparece en una forma irregular es: Dificultad: .32	*pudo N=524	corrieron N=268	comido N=416	se preocupa N=301		N=101

**Tabla 2.** Ítems representativos de la prueba y los patrones de respuestas dadas.

Para responder a la segunda pregunta, podemos concluir que la prueba parece ser inconsistente en cuanto a la forma y función de lo que solicita al informante y en cuanto a la forma en que estimula los conocimientos lingüísticos de los HLH. Algunos ítems representan información tan comúnmente conocida que sus respuestas son demasiado transparentes. Otros de los ítems incluyen elementos que compiten por la atención del examinado y resultan en engañarlo y desviarlo de la respuesta correcta. Finalmente, otros ítems piden que los HLH hagan caso omiso a las intuiciones lingüísticas que han desarrollado hasta el momento y que mejor intenten adivinar las reglas de la norma culta del idioma a la cual no han sido expuestos consistentemente. Entonces, en cuanto a la segunda consideración de validez, estos puntos levantan dudas nuevas sobre la prueba de ubicación.

En la tercera pregunta que propusimos, buscamos evidencia de una progresión de habilidad lingüística exhibida por los ítems. La información ofrecida en las Tablas 1 y 2 y la discusión anterior nos ayudan a responder a esta cuestión. En la Tabla 1 vemos que en todas las categorías de las 90 preguntas, la dificultad del ítem varía mucho de la proporción óptima esperada. Esta gran variabilidad nos lleva a cuestionar la eficacia de los ítems en su totalidad. Cuando examinamos los ítems a fondo para encontrar la raíz de su dificultad, como los ejemplos expuestos en la Tabla 2 y en la discusión anterior, encontramos que muchas de las preguntas que se clasifican como difíciles debido a problemas de la formación de la misma. Es decir, que la dificultad de la pregunta no se debe consistentemente a un incremento de complejidad, sino en varios casos a la ambigüedad o de la raíz o las alternativas ofrecidas en las respuestas. Por lo tanto, la “dificultad” de las preguntas se revela más como un artefacto de sus deficiencias que por una diferencia en el dominio de conceptos lingüísticos por parte de los examinados. Entonces, en cuanto a la tercera pregunta, la validez de la prueba sigue estando en duda.

La cuarta pregunta que propusimos no exige examinar la relación entre lo que examinamos en la prueba y las habilidades necesarias para los cursos hacia los que estamos guiando a los estudiantes. Los objetivos curriculares de los cursos en que se pueden ubicar los estudiantes por medio de la prueba se encuentran en la Tabla 3; incluye las descripciones abreviadas que aparecen en el catálogo escolar (NMSU 2008). Vemos que en las descripciones curriculares, el énfasis en el desarrollo de la lectoescritura no empieza de manera formal hasta el segundo curso de la secuencia, ya que reconoce que las destrezas de los estudiantes residen principalmente en la oralidad. El estudio de los elementos formales de la gramática no se inicia hasta el tercer curso de la secuencia y no llega a culminar hasta el último curso.

---

**SPAN 113 Beginning Spanish for Native Speakers**

Enfatiza la comprensión auditiva y el reconocimiento de vocabulario para la recuperación del idioma nativo. Actividades para aumentar confianza y estimular el desarrollo de destrezas orales.

**SPAN 213 Spanish for Native Speakers I**

Enfatiza el desarrollo de habilidades de leer en el idioma nativo. Actividades orales, escritura y vocabulario para mejorar el dominio del idioma. Únicamente para estudiantes hispanoparlantes.

**SPAN 214 Spanish for Native Speakers II**

Enfatiza la lectoescritura con actividades orales para estimular el desarrollo de habilidades. Conversaciones sobre áreas problemáticas de la gramática.

**SPAN 312 Grammar for Native Speakers of Spanish**

Para estudiantes que se han expuesto al español en casa o en la comunidad. Repaso de conceptos gramaticales y análisis tanto del español oral como escrito.

---

**Tabla 3.** Descripciones del catálogo estudiantil de los cursos para hablantes de español como lengua de herencia de New Mexico State University (traducción mía).

Podemos observar entonces, que tanto el formato de la prueba como el contenido contradicen estos objetivos. En primer lugar, la prueba se ofrece únicamente de forma escrita, aunque ahora se lee en pantalla en lugar de papel. De esta manera, obliga al examinado a organizar y transformar sus conocimientos, adquiridos oralmente y en contextos auténticos, a un formato textual. Los HLH más vulnerables son los HLH para quienes sus habilidades residen principalmente en sus capacidades receptivas, es decir, que solo entienden lo que oyen en español, pero casi no lo leen ni lo escriben ni lo hablan. La prueba no puede detectar adecuadamente las habilidades de estos estudiantes y guiarlos a la clase de 113. Los estudiantes elegibles para 213 tampoco se van a ubicar consistentemente ya que en esa clase es donde apenas se concentra en el desarrollo de habilidades de la lectoescritura. Por lo tanto el examen va a exhibir un sesgo hacia los cursos más avanzados. Los datos de la ubicación recomendada por parte de la prueba reafirman estas sospechas. De los 1610 identificados como HLH que presentaron el examen, 324 de los examinados fueron recomendados para la clase de 113 y otros 296 recibieron recomendación para la clase de 213. En cambio, 798 de los examinados fueron dirigidos a la clase de 214 y 192 a la clase de 312. Un análisis de chi-cuadrada, examinando la bondad de ajuste de estas cifras,  $\chi^2(3, N = 1610) = 542.2$ ,  $p < .0001$ , confirma el sesgo hacia los cursos más avanzados donde las habilidades de lectoescritura ya están establecidas.

En fin, no está claro lo que manifiestan los ítems de la prueba sobre las capacidades lingüísticas generales de los examinados, ni de cómo predicen el éxito del examinado en la clase recomendada. Lo que parece más evidente es que no estamos sondeando adecuadamente ni la competencia lingüística ni la competencia comunicativa de los HLH, ni estamos estableciendo los criterios mínimos requeridos para su éxito en la clase. Al contrario, mientras los cursos pretenden validar, avivar, y ampliar los conocimientos lingüísticos de los HLH, la prueba parece querer acentuar las faltas o lapsos de aprendizaje que no han experimentado todavía. De esta manera, en respuesta a la cuarta pregunta, no encontramos evidencia convincente de que la prueba cumpla con su función para nuestros fines de ubicación con los HLH de NMSU.

Por lo tanto, examinando los pilares de la validez sugeridos por los autores de los *Standards* (1999), no encontramos evidencia para apoyar ninguna de las cuatro consideraciones. En primer lugar, el contenido no es adecuado para averiguar la competencia lingüística de esta población de estudiantes. Segundo, hay inconsistencias con los ítems en cuanto a su nivel de dificultad y en cuanto al formato en que están escritos que llevan a los estudiantes no a activar sus habilidades lingüísticas, sino abandonarlas y adoptar otras estrategias. En tercer lugar, hemos visto que estos problemas influyen en el grado de dificultad anotado para los ítems; la dificultad se debe más a la confusión asociada con las preguntas, que por una mayor complejidad de contenido. Finalmente, no hallamos evidencia de un nexo entre lo que se les pide a los examinados en la prueba y en los objetivos de los diferentes niveles curriculares. Los varios problemas señalados con la forma en que se administra la prueba solo agravan estas dificultades.

Si consideramos la última preocupación sugerida por los autores de los *Standards* (1999), y especulamos sobre las consecuencias por tomar o administrar esta prueba, notamos que sí tiene un impacto negativo sobre los HLH. En primer lugar, es poco ético exigir a los estudiantes que tomen una prueba de ubicación cuyo resultado no tiene valor explicativo. La situación se empeora cuando consideramos la diferencia de lo que exigimos a los estudiantes no nativos con los de español como lengua de herencia. Los no nativos toman un examen compuesto de las 50 preguntas más básicas y, según el resultado, pueden exentarse de cursos hasta el segundo curso intermedio y recibir una recomendación para entrar al curso de gramática. Por contraste, los estudiantes identificados como hablantes de herencia tienen que presentar un examen de casi el doble de ítems, de un contenido más variado y más difícil, y sacar un puntaje mucho mayor para recibir la misma dispensación.

## 5. Conclusiones

Aunque en muchos casos utilizamos pruebas por rutina y sin inspeccionarlas, los estándares y las costumbres de la docencia y de las prácticas de evaluación nos obligan a realizar exámenes críticos regulares de los instrumentos que empleamos para verificar su valor continuo en cumplir su función. En este caso, la presente investigación crítica de la prueba de ubicación con referencia particular a los hablantes de español como lengua de herencia ha revelado problemas graves en cuanto a su contenido,

el formato de sus preguntas, la presentación de los ítems, y el uso y distribución de los resultados. En fin, se encontró que la prueba no es un instrumento válido ni para revelar las habilidades lingüísticas de los estudiantes HLH de nuestra región, ni para dirigirlos adecuadamente a los cursos de lengua de nuestro programa diseñados para ampliar y fortalecer sus destrezas.

Para remediar estas faltas y responder de manera ética a las necesidades de esta población de estudiantes que servimos, estamos en las primeras fases de crear una nueva prueba siguiendo las pautas de los *Standards* (1999) que recomiendan que se documenten las especificaciones, la planeación, y los métodos usados en el desarrollo de la prueba. Esta prueba será separada e independiente a la que se usa para los no nativos. Reflejará tanto las habilidades de la población de HLH, como las necesidades curriculares de nuestro programa. Incluirá ítems auténticos que se presentan en un orden lógico, y en formatos diversos, incluyendo formatos audiovisuales. Habrá un número de ítems suficiente para distinguir la competencia del estudiante, pero a una cantidad y grado de dificultad comparable con la prueba de los no nativos. Verificaremos que la prueba nos ayuda a colocar adecuadamente a los estudiantes HLH en los cursos iguales a su capacidad y seguiremos evaluando la prueba regularmente para confirmar su utilidad. De esa manera, la prueba se convierte en un instrumento válido y fiable que rinde beneficios tanto para los estudiantes HLH de nuestra región como para el programa que se dedica a servirles.

## Referencias

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Bachman, Lyle F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, Lyle F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press.
- Bills, Garland D., Hernandez Chavez, Eduardo, & Hudson, Alan. (1995). The geography of language shift: distance from the Mexican border and Spanish language claiming in the Southwestern U. S. *International Journal of the Sociology of Language*, 114, 9-27.
- Haladyna, Thomas, Downing, Steven. & Rodríguez, Michael. A Review of Multiple-Choice Item-Writing Guidelines for Classroom Assessment, *Applied Measurement in Education*, 15(3), 309–334.
- Hughes, Arthur. (2003). *Testing for Language Teachers*. Cambridge: Cambridge University Press.
- Kehoe, Jerard. (1995a). Writing multiple-choice test items. *Practical Assessment, Research & Evaluation*, 4(9). Retrieved November 10, 2005, from <http://PAREonline.net/getvn.asp?v=4&n=9>.
- Kehoe, Jerard. (1995b). Basic item analysis for multiple-choice tests. *Practical Assessment, Research & Evaluation*, 4(10). Retrieved November 10, 2005, from <http://PAREonline.net/getvn.asp?v=4&n=10>.
- MacGregor-Mendoza, Patricia. (2005). "El Desplazamiento Intergeneracional del Español en los Estados Unidos: Una aproximación". In L. A. Ortiz López y M. Lacorte (Eds.) *Contactos y contextos lingüísticos: El español en los Estados Unidos y en contacto con otras lenguas*, pp. 287-300. Madrid: Lingüística Iberoamericana.
- MacGregor-Mendoza, Patricia. (2001 [2000]). "Aquí no se habla español: Stories of linguistic repression in Southwest schools." *Bilingual Research Journal*, 24 (4): 333-345.
- MacGregor-Mendoza, Patricia. (1998). "Language and the Bilingual Teacher: Use, Attitudes, Roles", *Southwest Journal of Linguistics*, 17 (2), 83-99.
- New Mexico State University. (2008). 2008-2009 Undergraduate catalog. Las Cruces, NM: New Mexico State University.
- New Mexico State University. (2010). Quick Facts 2009-2010. Retrieved January 28, 2010, from <http://www.nmsu.edu/Research/iresearch/QuickFacts/Factsheets.html>.
- Norman, Laura M., Parcher, Jean W., & Lam, Alven H. (2004). Monitoring colonias along the United States-Mexico border: U.S. Geological Survey Fact Sheet 2004-3070. Retrieved January 28, 2010, from <http://geography.wr.usgs.gov/pubsAlpha.html#n>.
- Pease-Alvarez, Lucinda. (1993). Moving In and Out of Bilingualism: Investigating Native Language Maintenance and Shift in Mexican-Descent Children *NCRCDSSL Research Reports*.
- Piontek, Mary. E. (2008). Best Practices for Designing and Grading Exams. *CRLT Occasional Papers*, (24). Retrieved January 15, 2010, from [http://www.crlt.umich.edu/publinks/CRLT\\_no24.pdf](http://www.crlt.umich.edu/publinks/CRLT_no24.pdf).
- Rudner, Lawrence M. (1994). Questions to ask when evaluating tests. *Practical Assessment, Research & Evaluation*, 4(2). Retrieved November 10, 2005, from <http://PAREonline.net/getvn.asp?v=4&n=2>.

- Teschner, Richard V. (1983). Spanish placement for native speakers, nonnative speakers, and others. *ADFL Bulletin*, 14(3), 37-42.
- Teschner, Richard V. (1990). Spanish Speakers Semi- and Residually Native: After the Placement Test Is over. *Hispania*, 73, 816-822.
- The University of Texas at El Paso. (2008). The University of Texas at El Paso Facts 2007-2008 Retrieved January 28, 2010, from <http://admin.utep.edu/Default.aspx?tabid=50864>.
- U.S. Department of Transportation, (2009). U.S. Border Crossings/Entries by State/Port and Month/Year Sorted by Year (Ascending). Research and Innovative Technology Administration, Bureau of Transportation Statistics, Border Crossing/Entry Data based on data from U.S. Department of Homeland Security, Customs and Border Protection, OMR database. <http://www.transtats.bts.gov/BorderCrossing.aspx>.
- Weir, Cyril J. (2005). *Language Testing and Validation: An Evidence-Based Approach*. Houndsmills: Palgrave Macmillan.

# Selected Proceedings of the 13th Hispanic Linguistics Symposium

edited by Luis A. Ortiz-López

Cascadilla Proceedings Project Somerville, MA 2011

## Copyright information

Selected Proceedings of the 13th Hispanic Linguistics Symposium  
© 2011 Cascadilla Proceedings Project, Somerville, MA. All rights reserved

ISBN 978-1-57473-442-3 library binding

A copyright notice for each paper is located at the bottom of the first page of the paper.  
Reprints for course packs can be authorized by Cascadilla Proceedings Project.

## Ordering information

Orders for the library binding edition are handled by Cascadilla Press.  
To place an order, go to [www.lingref.com](http://www.lingref.com) or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA  
phone: 1-617-776-2370, fax: 1-617-776-2271, [sales@cascadilla.com](mailto:sales@cascadilla.com)

## Web access and citation information

This entire proceedings can also be viewed on the web at [www.lingref.com](http://www.lingref.com). Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

MacGregor-Mendoza, Patricia. 2011. Probando, probando...analizando la utilidad de un examen de ubicación para hablantes de español como lengua de herencia. In *Selected Proceedings of the 13th Hispanic Linguistics Symposium*, ed. Luis A. Ortiz-López, 150-160. Somerville, MA: Cascadilla Proceedings Project.  
[www.lingref.com](http://www.lingref.com), document #2483.