

## **Presentación**

Buenos días.

En primer lugar, y antes de comenzar con la defensa en sí, quisiera aprovechar esta oportunidad para agradecer a los miembros del tribunal su presencia aquí hoy. A la vez, también quisiera hacer público mis disculpas por las molestias que hayan sufrido a raíz de los cambios de fecha, y aplazamientos que han ocurrido. En los agradecimientos que vds habrán leído al principio de esta tesis, afirmo que había tardado mucho - demasiado - en terminarla. Algunas veces me ha parecido que íbamos a tardar aún más en llegar al día de hoy, en el que presento esta defensa de la tesis. ¿Será otra manifestación más de lo complicado que resulta trabajar e investigar en un mundo un tanto “desordenado”?

Dicho sea de paso, una ventaja de tanta demora - desde mi punto de vista personal - ha sido el hecho de que al igual que ocurre con las traducciones, cuando el texto “reposa” un tiempo, el traductor, o el autor, lo retoma de otra manera. Ahora, ya en el año 2.000, he convivido con esta tesis de forma más o menos íntima, durante muchos años pero después de estos últimos meses de “reposo” pienso poder defenderla de manera más objetiva.

Así que,— si les parece bien — comenzaré con un breve esbozo de lo que será esta defensa.

## ***Posición***

En primer lugar, para situarnos en el contexto actual describiremos, de la manera más concisa posible, los resultados de nuestro trabajo. Y para ello, seguiremos las líneas estructurales del diseño de la investigación. Ésta se divide en cuatro partes:

1. El estudio longitudinal,
2. El aspecto transversal,
3. La fiabilidad estadística,
4. La eficacia en la preparación de los instrumentos.

No es nuestra intención entrar aquí en cuestiones puntuales sobre las distintas pruebas que hemos confeccionado. Pero, por supuesto, intentaremos responder a cualquier observación que quieran hacer más adelante.

## ***El problema***

Una vez terminada esta puesta en escena, ubicaremos nuestro estudio en la situación actual, ya que las pruebas iniciales de aptitud volverán a aparecer en los debates de los departamentos y centros a lo largo del próximo cuatrimestre, y el tema del posible uso de una prueba única en todos los centros volvió a surgir en la reunión de centros que tuvo lugar en Salamanca hace poco.

## ***Posibilidades***

Hablaremos, pues, de algunas de las alternativas entre las que nosotros, aquí en Granada, y cada uno de uds. en sus respectivas universidades, podrían elegir, y propondremos dos campos de investigación en los que pensamos que nuestros centros podrían colaborar para profundizar aún más en este tema.

## Posición

Entonces, con su permiso, empecemos con una descripción de los cuatro vertientes principales de la investigación: el estudio longitudinal, el estudio transversal, la fiabilidad, y la eficacia en la preparación de los instrumentos.

1. Nuestro estudio comienza con la hipótesis de que las destrezas lectoras en la lengua A y la lengua B, constituyen un instrumento para medir la aptitud del individuo para la traducción. En el diseño de nuestro estudio, en el que aplicamos el modelo de Carroll (1981), averiguamos la validez predictiva de nuestras subpruebas, calculando las correlaciones pertinentes entre éstas y los notas de Traducción General A-B y B-A. De esta manera buscamos demostrar que las subpruebas reúnen la condición número 6 del modelo de investigación.

En su propuesta, Carroll no especifica ningún valor para esta correlación, solo propone que sea significativamente superior al cero. En cambio, nosotros fijamos una correlación igual a, or mayor de 0,65 como meta. En estas dos primeras transparencias que les voy a enseñar se ven claramente los niveles de correlación de las dos cohortes.

Si me permiten un breve inciso. Cada uno de uds. tiene copias de las transparencias entre los papeles que les he dejado encima de la mesa. En todas las transparencias el azul claro representa los datos de la Cohorte A, y el azul más oscuro, la cohorte B. Las columnas de color continuo presentan los datos estadísticamente significativos, y las de rombos, los datos no significativos. Las columnas de color morado representan los valores meta.

# **TRANSPARENCIA 1**

## **VALIDEZ PREDICTIVA – SUBPRUEBAS EN LENGUA B**

Estas correlaciones indican la validez predictiva de las subpruebas en Lengua B. No incluimos las de Lengua A porque eran más bajas, incluso negativas.

Ninguno de nuestros instrumentos alcanzó el nivel deseado. No obstante, en los resultados de la Cohorte A, estas correlaciones sí llegaron a niveles significativos.

Si tenemos en cuenta el hecho de que se modificó la especificación de la prueba entre la primera cohorte y la segunda, y de que existieron ciertas anomalías en la administración del conjunto de las pruebas a la Cohorte B, hoy nosotros afirmamos que la validez predictiva está probada. En el texto de nuestra tesis, nos limitamos a constatar que se podía rechazar la hipótesis nula de que no existe relación ninguna entre las destrezas lectoras en Lengua B y la Traducción General A-B y B-A. Hoy, pensamos que quizás deberíamos aceptar la hipótesis de que sí existe tal relación.

La dimensión longitudinal de este estudio consta de dos años, y evidentemente aquí tenemos uno de los grandes problemas de cualquier estudio longitudinal que es el hecho de que en dos años pueden introducirse muchas variables que pueden influir en el proceso y distorsionar los resultados. También pueden desaparecer muchos de los componentes del grupo inicial. Cuando nos encontramos con resultados estadísticamente significativos, a pesar de estos factores, pensamos que estos datos son realmente sólidos y nos reafirmamos en nuestra conclusión de que la validez predictiva de las subpruebas en Lengua B queda demostrada.



Además de la validez predictiva, teníamos la necesidad de confirmar el cumplimiento de la primera condición de Carroll, a través de la validez de respuesta de las pruebas. Ésta nos indica hasta qué punto las respuestas de los candidatos son coherentes con la prueba. Nosotros fijamos como mecanismo para calcular la validez de respuesta el error típico de medida. Este error se calcula una vez que se sabe la fiabilidad de la prueba.

En la literatura que hemos consultado nunca aparece ninguna recomendación acerca del valor aceptable de esta medida, por lo cual nosotros establecimos un valor igual a, o menor de 1,0 como meta.

## TRANSPARENCIA 2

### VALIDEZ DE RESPUESTA – ERROR TÍPICO DE MEDIDA ( $S_E$ )

Si ahora nos fijamos en la siguiente transparencia, veremos que, este valor no se ha podido alcanzar en ninguna de las cuatro subpruebas. El rango de valores pasa del 1,62 al 1,99. ¿Podemos, entonces, afirmar que también se ha cumplido la primera condición? Creemos que sí. Al comprobar que sí existe una correlación significativa entre las subpruebas en Lengua B y las notas de Traducción General, a pesar de la baja fiabilidad de aquellas, pensamos que al introducir las mejoras que más adelante señalaremos, el nivel de fiabilidad sería aceptable.

2. Si el objetivo primordial de unas pruebas de aptitud es predecir el comportamiento futuro de unos individuos en una situación específica, no nos podemos olvidar de la realidad del contexto en el que se realizan. El primer objetivo de todos los que trabajamos en la llamada *industria de los exámenes* tiene que ser confirmar que “el examen” sea realmente necesario. Si existen otros instrumentos que se pueden aplicar, o que ya se aplican, y que surten los mismos efectos, ¿por qué otro examen más? Los estudios de Bossé-Andrieu (1981) en el campo de la traducción, y toda la literatura sobre aptitud en lengua extranjera, indican que existen relaciones claras entre instrumentos que miden la inteligencia general, el acceso a la universidad, o los resultados de la enseñanza secundaria, y el comportamiento futuro en traducción y/o en el aprendizaje de una lengua B. Nuestra pregunta, pues, tenía que ser ¿Se justifica el uso de una prueba de aptitud inicial además de estas otras pruebas?

La validez predictiva de las notas de Selectividad, y de las de Educación secundaria contrastadas con las de Traducción General es inferior a la de nuestras subpruebas, y carece de significado estadística, como demuestran las siguientes transparencias:

### **TRANSPARENCIA 3**

**VALIDEZ PREDICTIVA – SELECTIVIDAD Y TRAD.GRAL.**

### **TRANSPARENCIA 5**

**VALIDEZ PREDICTIVA – SECUNDARIA Y TRAD.GRAL.**

La validez concurrente de las mismas notas, contrastadas ahora con las de nuestras subpruebas indica un solapamiento entre los conocimientos y destrezas examinados que, en el caso de la Cohorte A, sí es significativo.

### **TRANSPARENCIA 4**

**VALIDEZ CONCURRENTE – SELECTIVIDAD Y SUBPRUEBAS**

### **TRANSPARENCIA 6**

**VALIDEZ CONCURRENTE – SECUNDARIA Y SUBPRUEBAS.**

Pero lo que es mucho más interesante son las correlaciones entre la Oxford Placement Test, y Traducción General, y la OPT y nuestras subpruebas:

## **TRANSPARENCIA 7**

### **VALIDEZ PREDICTIVA – OPT Y TRAD. GRAL.**

Aquí se ve claramente que la OPT es el mejor instrumento para predecir comportamiento en traducción. Mejor, pero imposible de usar ya que está a la venta en las librerías.

Y esta otra transparencia, se ve que el solapamiento entre las destrezas y conocimientos de las pruebas es grande, pero no completa.

## **TRANSPARENCIA 8**

### **VALIDEZ CRITERIAL – OPT Y SUBPRUEBAS.**

Si ahora quisieramos ordenar los cuatro instrumentos, veríamos que la OPT es el que mejor predice aptitud para traducir, después vienen nuestras pruebas, seguido por las notas de Selectividad, y en último lugar las notas de Secundaria.

3. La baja fiabilidad estadística de los instrumentos que hemos utilizado para elaborar nuestro estudio es uno de los elementos que ya hemos comentado. Afecta, como recordarán, a la validez de respuesta debido a que es la base de la fórmula que se utiliza para calcular el error típico de medida. Ahora bien, la siguiente transparencia demuestra la manera en la que la fiabilidad de nuestras subpruebas podría mejorarse.

## **TRANSPARENCIA 9**

### **FIABILIDAD – POR MITADES Y PREDICTIVA.**

Las columnas representan de las subpruebas que analizamos, y la parte inferior es la fiabilidad real. La parte superior, en verde, equivale al aumento de fiabilidad que conseguiríamos al aumentar el número de ítems incluyendo preguntas homogéneas con las que actualmente conforman las pruebas. Como ven, para conseguir mayor fiabilidad, las pruebas tendrían que ser bastante más largas.

4. Por último, vamos a comentar los aspectos que más conciernen la eficacia en la preparación de las subpruebas.

En este trabajo, utilizamos tareas de examen como mecanismos para medir los constructos de destrezas lectoras que habíamos designados indicativos de aptitud para traducir. Las cinco tareas que hemos utilizado han sido:

Para la Cohorte A

Actividades de relacionar

Ítems Verdadero/Falso/No se sabe

Ítems de selección múltiple con cinco opciones;

y para la Cohorte B

Ítems de selección múltiple con cuatro opciones

Actividades de identificar errores

Actividades de corregir errores

Pero el enfoque de cada ítem ha sido una de las destrezas lectoras descritas por Munby (1978) y las combinaciones de éstas tareas con las destrezas han sido el objetivo fundamental de nuestro análisis de los valores de índice de facilidad e índice de discriminación. Nosotros buscamos relacionar tareas y destrezas de manera que pudiéramos especificar en el futuro cuales serían las combinaciones más eficaces, y así facilitar el proceso de elaboración de las pruebas.

Nuestro procedimiento para juzgar el comportamiento de los ítems ha sido el de contrastar los valores de los dos índices con nuestras metas: entre 0,3 y 0,7, para el índice de facilidad, e igual a, o mayor de 0,3 para el índice de discriminación. Con más frecuencia los ítems de selección múltiple han producido los resultados más acertados, y éstos se han encontrado en combinación con tres destrezas:

Deducir el significado e uso del léxico desconocido en función del contexto en el que aparece

Comprender información explícita en el texto

Comprender información implícita en el texto, por inferencia

Los resultados de las otras destrezas, en combinación con cualquiera de las clases de tarea, han resultado poco convincentes por inconsistentes, aunque sí han producido valores dentro de los parámetros establecidos.



## **El problema**

Pues bien, con esto terminamos nuestro repaso de los resultados y pasamos a la realidad. La cuestión de las pruebas sigue en pie. A raíz de nuestro estudio, pensamos que queda demostrada empíricamente la necesidad de realizar una prueba de aptitud inicial con el fin de asegurar una selección de los alumnos más capacitados para llevar a cabo los estudios en traducción con éxito. Nuestros resultados indican que la relación entre destrezas de lectura en Lengua B y traducción existe, pero que intervienen otros elementos, y que la realidad es más compleja que la que nuestro planteamiento inicial preveía.

Ahora, vamos a analizar cuatro alternativas.

## **Posibilidades**

La primera opción es la que acabamos de descalificar, la de no emplear ninguna prueba. La mencionamos porque sabemos que es la opción empleada en algunos centros — entendemos que en la Universidad de Málaga, por ejemplo, se sigue sin usar ninguna prueba — y porque es una opción que, de elegirla en Granada, tendría unas consecuencias claras.

Si de hecho, en las dos cohortes de nuestro estudio la nota media de Selectividad de los que entraron fue de 7,92 y 8,12, respectivamente, la eliminación de la prueba seguramente aumentaría esta media: entrarían alumnos mejor preparados. Pero, la contrapartida de este aumento de nivel académico sería una bajada en el nivel medio de Lengua B inglés, y el peso de esta situación la llevarían los profesores que imparten la asignatura de Lengua B. Ellos se encontrarían con grupos más heterogéneos, y con unas diferencias de nivel mayores.

La actitud en algunos países francófonos sería la de aceptar esta situación ya que “la lengua B, se puede aprender”.

La segunda alternativa sería la de mantener la prueba integrada, basada en la redacción de un resumen de un texto oral. Esta es la que en los últimos años se ha vuelto a emplear en Granada, la que se usa en la Universidad de Alicante, y en la Autónoma de Barcelona, entre otros. Estas pruebas requieren una inversión en tiempo y preparación para asegurarse de que los criterios que se emplean en la corrección de las pruebas son coherentes. La fiabilidad inter-evaluador, e intra-evaluador son los grandes retos de estas pruebas. Aquí me remito a los trabajos de la Profesora Dra Olivia Fox, de la Universidad Autónoma de Barcelona, que está trabajando en este campo, pero que todavía no posee resultados concluyentes. El debate en Barcelona, se ha pasado al campo políticas, que es inevitable en el mundo de los exámenes.

La tercera alternativa, y la que seguramente es la más completa, es muy similar a la de la Universidad Pontificia de Comillas. Emplear una batería de pruebas, analíticas y sintéticas, de destrezas receptoras y productoras, nos proveería de un amplio abanico de datos acerca de los conocimientos de los candidatos. La dificultad de esta opción está en la enorme inversión que supondría, y la dificultad que las universidades públicas tendrían en llevarlo a cabo.

Por último, llegamos a la opción que nos parece la más adecuada dentro de las limitaciones que nosotros tenemos aquí en la Universidad de Granada. Esta sería la de emplear una batería de subpruebas, incluidas pruebas similares a las nuestras, en combinación con pruebas objetivas de producción escrita, parecidas a las que aparecen en la OPT. Creemos que así, llegaríamos a diseñar unas pruebas adecuadas a la necesidad y a las circunstancias.

Y después de comentar estas cuatro alternativas, para terminar, quisieramos dibujar los dos campos de investigación en los que pensamos que nuestros centros deberían de colaborar para profundizar en la cuestión de aptitud para traducir.

Nuestra investigación aporta datos sobre la situación con referencia a la lengua B inglés en la Universidad de Granada. Pero somos conscientes de que no es suficiente. Pensamos que interesaría recoger datos comparativos de la situación de las otras lenguas B, y de los otros centros. Datos que proporcionarían conocimiento de la situación actual en cuanto a los alumnos que entrán en traducción, y en cuanto a los niveles de conocimientos y destrezas adquiridos en el primer ciclo de estudios. Para ello, pensamos que los trabajos que han realizado las asociaciones que componen la ALTE servirían como base excelente para comenzar este trabajo.

Datos también, deberíamos de reunir — y aquí me remito a la valiosa tesis doctoral del Profesor Dr Waddington — sobre los criterios que se emplean a la hora de corregir las traducciones en Traducción General.

Un estudio comparativo desde el punto de partida, y hasta el punto de llegada de nuestros alumnos nos serviría como fundamento para diseñar unas pruebas de aptitud más específicas.

Señores, hemos repasado los puntos más importantes de este trabajo, y hemos intentado ubicarla en el contexto actual de los estudios en Traducción en España. Les hemos ofrecido una vista parcial de alternativas que se podrían emplear, y por último hemos querido sugerir dos campos de investigación en los que muchos podríamos colaborar.

Sinceramente, les agradezco su atención.