

## Desarrollo de un test de nivel de comprensión pragmática para estudiantes sinohablantes de español como lengua extranjera

*Development of a test to assess the pragmatic level of comprehension for speakers of Spanish as a foreign language*

**M. Isabel Gibert-Escofet**

Universitat Rovira i Virgili  
España

**Mar Gutiérrez-Colón Plana**

Universitat Rovira i Virgili  
España

**Pedro-Manuel Medina Garcia**

Departament de Justícia de la Generalitat de Catalunya  
España

ONOMÁZEIN 49 (septiembre de 2020): 115-139

DOI: [10.7764/onomazein.49.07](https://orcid.org/0000-0003-3158-9658)

ISSN: 0718-5758



**M. Isabel Gibert-Escofet:** Departament de Filologies Romanicas, Campus Catalunya, Universitat Rovira i Virgili, España. <https://orcid.org/0000-0003-3158-9658>. | E-mail: [isabel.gibert@urv.cat](mailto:isabel.gibert@urv.cat)

**Mar Gutiérrez-Colón Plana:** Department of English and German Studies, Campus Catalunya, Universitat Rovira i Virgili, España. <https://orcid.org/0000-0001-8479-4933>. | E-mail: [mar.gutierrezcolon@urv.cat](mailto:mar.gutierrezcolon@urv.cat)

**Pedro-Manuel Medina Garcia:** Departament de Justícia de la Generalitat de Catalunya, España. | E-mail: [pmmg8@yahoo.es](mailto:pmmg8@yahoo.es)

Fecha de recepción: agosto de 2018

Fecha de aceptación: agosto de 2019

## Resumen

El objetivo principal de este estudio es la creación de un test de nivel de comprensión pragmática para estudiantes sinohablantes de español como lengua extranjera. Los resultados de este estudio muestran un test estandarizado que cumple con los requisitos de utilidad (autenticidad, interactividad, impacto, practicabilidad y valoración y puntuación) propuestos por Bachman y Palmer (1996) para las pruebas de lengua y los requisitos estadísticos psicométricos de fiabilidad y validez.

**Palabras clave:** pragmática; adquisición de segundas lenguas; español como lengua extranjera; aprendices sinohablantes; evaluación de la pragmática.

## Abstract

The main objective of this study is the creation of a test to assess the pragmatic level of comprehension for students who study Spanish as a foreign language. The results of this study show a standardized test that complies with the requirements of utility (authenticity, interactivity, impact, practicability and assessment and scoring) proposed by Bachman and Palmer (1996) for language tests and psychometric statistical requirements of reliability and validity.

**Keywords:** pragmatics; second language learning; Spanish as a second language; Chinese language learners; pragmatics assessment.

## 1. Introducción

El componente pragmático forma parte de la competencia comunicativa de una lengua y su dominio hace posible el uso adecuado de la misma en cada situación comunicativa específica (Canale y Swain, 1980; Bachman, 1990) y un test de pragmática como tal es una herramienta que resulta útil para mejorar la docencia de grupos meta concretos y la investigación en la adquisición del español como lengua extranjera (ELE) o segunda lengua (EL2).

En la actualidad no existe ningún test de nivel para medir la comprensión pragmática del español y tampoco existe, pues, ninguno que sea específico para estudiantes sinohablantes. La popularidad del español entre los estudiantes chinos universitarios ha incrementado de forma significativa su presencia en las universidades españolas. Los profesores de ELE han detectado una carencia de materiales para este perfil de alumnado. La creación de materiales que aborden de forma específica las necesidades que surgen de la gran distancia cultural y lingüística entre el chino y el español podrá llevarse a cabo con mayor precisión si conocemos el nivel de comprensión pragmática de los estudiantes.

El objetivo principal de este trabajo es crear un test de nivel de comprensión pragmática diseñado especialmente para los estudiantes chinos que se encuentran estudiando en una universidad española. Se trata de un test que siga los criterios de utilidad y fiabilidad y que esté estadísticamente validado para poder conocer su nivel de pragmática y, así, compararlo con su nivel de gramática (ya que no siempre se corresponde con su nivel de gramática).

En la figura 1 queda sistematizado todo el proceso seguido en la realización del test de comprensión pragmática para estudiantes sinohablantes que se explicará en los siguientes apartados.

### FIGURA 1

Proceso seguido en la realización del test de comprensión pragmática para estudiantes sinohablantes

ESTUDIO	OBJETIVOS
Estudio preliminar (generación ejemplar)	<ul style="list-style-type: none"> <li>– Obtener situaciones reales cotidianas en la vida de un estudiante universitario (conseguir situaciones auténticas) para la redacción de los ítems del test de nivel de comprensión pragmática.</li> <li>– Estudio etnográfico para conocer la percepción de nivel de dificultad de los distintos escenarios por parte de los estudiantes chinos.</li> <li>– Estudio etnográfico para recoger escenarios habituales entre los estudiantes universitarios chinos en España.</li> </ul>
<i>Recogida de datos</i>	
Encuesta 1	– Recoger datos sobre las situaciones en contexto universitario y su grado de dificultad (encuesta 1).
Encuesta 2	– Recoger datos sobre las situaciones en contexto universitario y su grado de dificultad (encuesta 2).

Encuesta 3	– Recoger datos sobre las situaciones en contexto universitario y su grado de dificultad (encuesta 3).
Encuesta 4	– Recoger datos sobre las situaciones en contexto universitario y su grado de dificultad (encuesta 4). – Obtener situaciones consideradas muy difíciles por los estudiantes.
Estudio principal (test de nivel de pragmática)	– Describir en detalle los componentes del diseño del test de nivel de comprensión pragmática. – Desarrollar el test de nivel de comprensión pragmática.
<i>Recogida de datos</i>	
Test piloto 1 (test de pragmática)	– Comprobar el grado de autenticidad que presentan los enunciados de los ítems según los hablantes nativos expertos (profesores de ELE).
Test piloto 2 (test de pragmática)	– Comprobar que los ítems no contienen ambigüedades y son comprensibles para los hablantes no nativos estudiantes chinos de ELE. – Recoger nuevos datos que sirvan para redactar distractores específicos para el grupo meta.
Test piloto 3 (test de nivel de gramática y test de nivel de pragmática)	– Verificar que las preguntas del test de nivel de gramática y del test de nivel de pragmática no contienen errores según los hablantes nativos.
Test de nivel de gramática y test de nivel de comprensión pragmática	– Conocer el nivel de gramática de los estudiantes chinos de ELE. – Conocer el nivel de comprensión pragmática de los estudiantes chinos de ELE.
Retest de nivel de comprensión pragmática (validación y fiabilidad)	– Comprobar la fiabilidad estadística del test en los estudiantes chinos de ELE.

Fruto de los enfoques comunicativos introducidos en los años 70, los test se han ido adaptando y se han centrado en medir sobre todo la competencia de los aprendientes desde una perspectiva funcional (Oller, 1979; Canale y Swain, 1980; Farhady, 1980; Canale, 1983; Bachman, 1990; Bachman y Palmer, 1996; Shohamy, 1996; Bachman y Palmer, 2010). Sin embargo, el test aquí propuesto va más allá de esta competencia funcional y tiene como objetivo separar la competencia de la comprensión pragmática del conocimiento y la competencia lingüística de los estudiantes.

Para diseñar y desarrollar el test de nivel de comprensión pragmática, se ha usado una metodología correlacional o psicométrica, mediante la cual se describen variables y se analizan sus relaciones a partir de una muestra representativa de sujetos. La información se recoge administrando a los sujetos que forman la muestra un test estandarizado que cumpla con los requisitos de utilidad propuestos por Bachman y Palmer (1996) y los requisitos estadísticos psicométricos de fiabilidad y validez. Es obvio, pues, que el desarrollo del test ha conllevado tomar

decisiones sobre ciertos aspectos: definir el constructo, seleccionar la muestra, cómo administrar el test, cómo codificar y analizar la información obtenida y qué conclusiones podemos extraer de los datos recogidos. Todo esto se expone y analiza en los apartados que siguen.

## 2. Marco teórico

Este estudio define en detalle los componentes del diseño del test que nos permitirá saber que la comprensión que los participantes tienen de la pragmática se acerca al máximo a la realidad y que la puntuación que se les va a otorgar será útil para los objetivos de los administradores.

Todas las decisiones que se han ido tomando para elaborar el test han tenido en cuenta las seis cualidades propuestas por Bachman y Palmer (1996: 17): (1) *fiabilidad*, es decir, la coherencia de medición; (2) *validez del constructo*, una interpretación demostrada con pruebas de que los resultados del test miden la competencia de la lengua que queremos medir; (3) *autenticidad*, que relaciona las características de las tareas propuestas en el test con las características de las tareas del uso de la lengua meta; (4) *interactividad*, una cualidad profundamente implicada en los constructos que queremos evaluar y en las corrientes actuales de enseñanza-aprendizaje de lenguas extranjeras; (5) *impacto*, relacionado con el uso del test y su influencia en la sociedad; y, por último, (6) *factibilidad*, que está relacionada con la implementación del mismo (un test factible es aquel que puede ser implementado con los recursos disponibles, es decir, cuyo diseño, desarrollo y uso no exigen más recursos de los que ya existen).

El test de nivel de comprensión pragmática sigue el marco propuesto por Hudson, Detmer y Brown (1992) y toma como ejemplo el diseñado por Roever (2005), así como las directrices marcadas por Bardovi-Harlig y Dörnyei (1998), Roever (2001) y Schauer (2009), aunque los ítems del test han sido diseñados específicamente para los estudiantes de español sinohablantes que se encuentran en situación de *study abroad* y no se adentra en el estudio de la producción.

## 3. Metodología

Antes de empezar a crear el test, se realizó un estudio etnográfico (que se corresponde con el estudio preliminar) partiendo del método de generación ejemplar, en el que se presentaban situaciones cotidianas de la vida universitaria española y los estudiantes chinos tenían que valorar en una escala del 1 al 10 el grado de dificultad que pensaban que conllevaban estas situaciones. Para ello se realizó un estudio etnográfico partiendo del método de generación ejemplar y que sirvió para detectar el grado de dificultad que se perciben en las situaciones más comunes de la vida universitaria de un alumno.

Para la identificación y selección del uso de la lengua meta y para poder decidir cuáles son las tareas apropiadas y las muestras de lengua para este test, nos hemos centrado en el aná-

lisis de necesidades del grupo meta al que va dirigido y el tipo de actividades identificadas como eficientes por anteriores estudios (Hudson, Detmer y Brown, 1992; Brown y Hudson, 2002; Roever, 2005; Schauer, 2009; Taguchi, 2008).

El análisis de necesidades se ha llevado a cabo siguiendo el procedimiento establecido por Bachman y Palmer (1996: 102): (1) la identificación que hacen los candidatos de las situaciones más relevantes en las tareas cotidianas que deben llevar a cabo en un contexto universitario de *study abroad*; (2) la recogida de información, la identificación y desarrollo de procedimientos para diseñar las tareas; (3) la recogida de información sobre el dominio y las tareas que se van a evaluar en un trabajo común con estudiantes que cumplan el perfil del grupo meta; (4) el análisis de las características de las tareas; y (5) la organización de las tareas con características similares por grupos. La obtención de los ítems se ha realizado a través del PCIC (Instituto Cervantes, 2007) y de las situaciones proporcionadas por el estudio preliminar. Además, se han consultado corpus de ELE (Val.Es.Co, GRIALE, CORELE) para ver si ayudan a redactar los ítems correctamente.

Para el diseño de las preguntas del test, se ha dividido el test en tres secciones: rutinas conversacionales, implicaturas y actos de habla (ver ilustraciones 1, 2 y 3 en los anexos). Como se trata de un test de nivel de comprensión pragmática, en los tres casos se diseñan actividades de elección múltiple con cuatro opciones por pregunta, como proponen Hudson, Detmer y Brown (1992, 1995) y, luego, lo aplica Roever (2005). Teniendo en cuenta a Brown (2001), que aconseja no elaborar test de lengua con más de 100 preguntas, y a Roever (2005), que elabora un test de 18 preguntas por sección, se fija el número de ítems del test de nivel de comprensión pragmática en 54.

Sin embargo, debido al procedimiento de desarrollo estadístico del test, es necesario aumentar el número de ítems por sección de 18 a 24 (72 para la totalidad del test), es decir, crear 6 ítems más por sección con el objetivo de poder eliminar después de pilotar el test las preguntas que no sean estadísticamente fiables. Esta técnica permite tener la opción de descartar aquellas preguntas que no se consideren válidas si se identifican como no significativas. Las preguntas no significativas son aquellas que proporcionan un número de errores o aciertos mayoritario o que se identifiquen como ambiguas. Para diseñar los ítems del test hay que tener en cuenta su autenticidad.

En el test resultante de este estudio, las preguntas que se han diseñado no están relacionadas con el contexto académico, sino con la *vida real o cotidiana* de un estudiante universitario.

### 3.1. Definición del constructo

El test se elabora sobre un constructo teórico no medible de manera directa; por ello se hace necesario el desarrollo del test que nos ayude a medirlo a través de evidencias empíricas y observables (Thorndike y Hagen, 1980). El constructo define lo que queremos medir con el test y un test solo será válido si validamos su constructo, es decir, si demostramos que mide aquello que efectivamente se quiere medir (Thorndike y Hagen, 1980).

Este test mide el nivel de comprensión pragmática que los estudiantes tienen de las rutinas conversacionales, la implicaturas y los actos de habla, componentes de los conceptos básicos sobre los que está elaborado, la pragmalingüística y sociopragmática. Entendemos por rutinas conversacionales aquel enunciado que sigue unas pautas lingüísticas y extralingüísticas (Gumperz y Hymes, 1972) que se repiten siempre en los mismos contextos de forma ritualizada. Su uso inadecuado puede causar conflictos, ya sea por descortesía, ya sea por desconcierto en las relaciones sociales. Siguiendo los criterios establecidos para las otras dos secciones del test, en esta sección también se ha escrito la batería de ítems correspondiente a las rutinas conversacionales considerando las características que marca Roever en su propia taxonomía (2005: 47). Estas condiciones son las siguientes: todos los personajes deben tener un nombre; en cada ítem el intercambio es entre un hombre y una mujer; hombre y mujer deben estar distribuidos de forma más o menos uniforme en cuanto a la elección que debe tomar el candidato; hay que definir la relación entre los interlocutores (amigos, compañeros, alumnos...); las situaciones tienen que resultar familiares a los candidatos.

Entendemos por implicaturas aquellos enunciados que emitimos cuando hablamos de forma indirecta, es decir, la información de los enunciados no se transmite de forma explícita.

Entendemos por actos de habla aquellas unidades básicas de comunicación que transmiten la intención del hablante; como tales, no pueden calificarse de correctos o incorrectos, ya que lo que pretenden no es afirmar que un enunciado es verdadero o falso, sino si es adecuado o inadecuado dependiendo del contexto en el que se produzca (Austin, 1976 [1962]). Para diseñar los ítems que componen la batería de preguntas de la sección de actos de habla del test, se siguen las condiciones que marca Roever (2005: 49) en su estudio, adaptadas en su mayoría de Johnston, Kasper y Ross (1998). Además, se toma la clasificación de Searle para buscar un equilibrio entre los distintos actos de habla. Sin embargo, en las situaciones presentadas en el test, el grado de imposición, en general, es bajo, ya que los intercambios suelen ser entre amigos o compañeros; los intercambios entre profesores o desconocidos son menos abundantes, ya que en la vida real de los estudiantes también es así. La distancia social y el poder diferencial siempre son bajos. Existe acuerdo con Yamashita (2008) en que es más apropiado incluir otros actos de habla como mandatos, promesas o agradecimiento, entre otros, para poder medir la pragmática de la interlengua de los candidatos.

### 3.2. Definición de las variables

A continuación se relacionan sus definiciones conceptuales y operacionales:

- a) variables independientes asignadas: sexo, edad, nivel de español, pareja española; tiempo de residencia en España o en un país hispanohablante; estudios universitarios;
- b) variable dependiente asignada: nivel de comprensión pragmática.

### 3.3. Desarrollo del test: operacionalización

La operacionalización del test implica la descripción del tipo de pregunta de la que va a constar y un borrador o proyecto (Bachman y Palmer, 1996: 90) en el que se describa su organización. En esta fase también hay que incluir el desarrollo y elaboración de los ítems, la creación de las instrucciones y enunciados, así como las especificaciones del método de corrección y la puntuación.

La operacionalización del test viene dada por estos tres principios:

Principle 1. It is the overall usefulness of the test that is to be maximized, rather than the individual qualities that affect usefulness. Principle 2. The individual test qualities cannot be evaluated independently, but must be evaluated in terms of their combined effect on the overall usefulness of the test. Principle 3. Test usefulness and the appropriate balance among the different qualities cannot be prescribed in general, but must be determined for each specific testing situation. (Bachman y Palmer, 1996: 18)

Este test va a ratificar que es posible medir el nivel de comprensión pragmática de forma fiable con un test de calidad para un grupo de aprendientes específico. El proceso seguido para el desarrollo del test ha sido laborioso y ha requerido de varias fases: (1) redacción de los ítems para el test piloto a partir de los escenarios obtenidos en el estudio preliminar; (2) administración del test con 24 ítems por sección para poder eliminar los ítems ambiguos y los poco auténticos; (3) revisión de los resultados del test piloto administrado a los nativos con el feedback de autenticidad que estos aportan; (4) redacción y eliminación de los ítems revisados por los expertos nativos; (5) redacción del cuestionario para pilotar entre estudiantes no nativos; (6) revisión del cuestionario con especial atención a los ítems que resulten ambiguos o que den como resultado un 100% de aciertos o errores; (7) redacción final del test con la eliminación de las preguntas conflictivas, modificando aquellas poco auténticas por propuestas hechas por los nativos e incluyendo propuestas de distractores proporcionados por los estudiantes chinos que forman parte de la administración piloto del test; (8) redacción de un test previo al test de comprensión pragmática para verificar el nivel de gramática de español de los estudiantes que participan en este estudio.

### 3.4. Desarrollo y descripción de los test piloto

El test está estructurado en tres secciones: rutinas conversacionales, implicaturas y actos de habla. Los distractores están basados todos en la adecuación pragmática, no tienen errores gramaticales, aunque sí formas no adecuadas en español. Para asegurar que las preguntas redactadas y las interacciones sean totalmente auténticas en cuanto a contexto y a lengua oral, se administran tres test piloto en tres fases distintas.

#### Fase 1 (test piloto 1): nativos

En una primera fase, el test piloto 1 se administra a expertos hablantes nativos (18 profesores de ELE). Su objetivo es comprobar el grado de autenticidad que presentan los enunciados de

los ítems ya que la norma del hablante nativo se reconoce como la adecuada (Kasper, 1995; North, 2000; Roever, 2000; McNamara y Roever, 2006).

El test presenta cuatro posibles respuestas para cada ítem más una respuesta abierta llamada "otra" en la que el participante puede añadir su propuesta de muestra de lengua si considera otra más real. De entre todas las opciones presentadas como posibles respuestas se pueden elegir dos, en cuyo caso se activa una pregunta dinámica en la que se debe responder cuál de ellas es la respuesta preferida. Después de estas dos preguntas, se introduce en el test un apartado en el que los nativos expertos deben dar un grado de autenticidad a cada una de las respuestas y, para ello, hay que marcar una opción en una escala de cuatro grados que son los siguientes: "Nada real", "Poco real", "Bastante real", "Real". Finalmente, se incluye un ítem en forma de pregunta abierta en el que los colaboradores nativos pueden añadir observaciones y comentarios sobre el ítem (ver la imagen 1).

## IMAGEN 1

Proceso de creación de las respuestas

**SECCIÓN 1: RUTINAS NATIVOS (USO DE LA LENGUA EN SITUACIONES CONCRETAS)**

**4.- PREGUNTA 1 - N1-S130-S38-C3-PCIC F 5.13.**

**INSTRUCCIONES:**

Vas a leer una serie de situaciones y tú vas a tener que decidir qué dirá la persona que se encuentra en esa situación. Cuando lo hayas decidido, clics en la opción correcta y luego en el botón "siguiente".

1. Elige la opción que te parezca más adecuada.
2. Si dudas entre dos, marca las dos.
3. Si has marcado dos respuestas, marca cuál es tu preferida.
4. Si se te ocurre una mejor puedes escribirla en el recuadro destinado a ello.

**\* Ana se toma un café en el bar y deja propina. El camarero le da las gracias.**  
**Camarero: Gracias.**  
**¿Qué crees que dirá Ana?**

a) "No, no."  
 b) "Para qué."  
 c) "Eso es todo."  
 d) "De nada."  
 Otra (por favor, especifica)

**\* He marcado dos respuestas, pero mi respuesta preferida es:**

a  b  c  d  Otra

**\* ¿Qué grado de veracidad le das a la situación?**

Nada real  Poco real  Bastante real  Real

**Te agradezco cualquier observación o comentario que quieras hacerme:**

Para redactar las interacciones y situaciones orales presentadas en el test se han consultado manuales, tenido en cuenta sugerencias de profesores expertos conocedores del contexto universitario y adaptado ejemplos del PCIC (Instituto Cervantes, 2007). Todas las opciones proporcionadas como posibles respuestas son gramaticalmente correctas, aunque solo una de ellas es pragmáticamente adecuada. Al redactar los distractores, es decir, las opciones ofrecidas como respuestas no adecuadas, se ha considerado el contexto y la pragmática china para que estos sean adecuados en la cultura china y poco adecuados en la española. Estos distractores también se han obtenido a partir de las respuestas no adecuadas y las respuestas sugeridas por estudiantes chinos que han formado parte del test piloto a hablantes no nativos.

### **Fase 2 (test piloto 2): no nativos**

El test creado en la fase 1 (test piloto 1) se administra a hablantes no nativos (12 estudiantes chinos de ELE con un nivel de español de entre B2 y C2). Su objetivo es, por una parte, comprobar que los ítems no contienen ambigüedades y son comprensibles y, por otra parte, recoger nuevos datos que sirvan para redactar distractores específicos para el grupo meta. El resultado es un test piloto 2.

### **Fase 3 (test piloto 3): test definitivo**

El test piloto 2 se administra a 18 hablantes nativos, estudiantes y población general para verificar que no existen fallos en la elaboración de las preguntas. En este test se elimina la posibilidad de poder elegir dos de las cuatro opciones propuestas como respuesta, el usuario está obligado a escoger solo una opción y se eliminan las preguntas sobre el grado de autenticidad y las preguntas abiertas para sugerencias. Se trata de un test mucho más asequible en términos de tiempo. Con las respuestas de estos dos colectivos constatamos que el test resultante podrá administrarse como test definitivo, ya que sus 54 ítems (18 por sección) no son susceptibles de ambigüedad. Todos los hablantes nativos identificaron como adecuadas las mismas respuestas. La estructura de este tercer test piloto es prácticamente igual que los dos anteriores, con la salvedad de los cambios introducidos para su uso final. Una vez analizados los resultados, revisado y modificado coincide con el test definitivo.

## **3.5. Descripción de la muestra de estudiantes participantes para el test final**

La selección de los estudiantes participantes en el estudio se hace por muestreo aleatorio probabilístico estratificado porque los sujetos deben tener distinto nivel de pragmática y este tipo de muestreo considera categorías típicas entre sí (estratos), con mucha homogeneidad (en este caso, la nacionalidad, el ser estudiantes universitarios, la lengua materna, etc.). De esta forma, todos los estratos de interés se ven representados en la muestra. Debido a la dificultad de acceso a la población, se trabaja también con un muestreo incidental, ya que con este tipo de muestreo el investigador puede seleccionar directamente los sujetos que formarán parte de la muestra.

Los estudios se diversifican en Diploma de Estudios de Lengua y Cultura Hispánica, Curso Universitario de Especialización en Lengua, Filología Hispánica, Máster en Enseñanza de ELE, Curso de Especialización de Empresa y Lengua Española, Curso de Especialización de Turismo y Lengua Española, Filología Francesa, Relaciones Laborales y Recursos Humanos, Ingeniería de Sistemas de Telecomunicación, Máster en Contabilidad y Auditoría Superior.

De la Universidad Rovira i Virgili (un 80% en el test) se eligieron estudiantes de tres cursos diferentes (Diploma de Estudios de Lengua y Cultura Hispánica; Curso Universitario de Especialización en Lengua (itinerario: lengua española y cultura hispánica); Máster en Enseñanza de Lenguas Extranjeras (itinerario: Español Lengua Extranjera); Curso de Especialización de Empresa y Lengua Española; y Doctorado); un 20% procedían de universidades distintas a la URV (20 de la Universidad Complutense de Madrid, 10 de la Universidad de Lérida, 6 de la Universidad Rey Juan Carlos, 4 de la Universidad de Barcelona, 2 de la Universidad Autónoma de Barcelona, 2 de la Universidad Politécnica de Madrid, 2 de la Universidad de Málaga, 2 de la Universidad de Zaragoza, 2 de la Universidad Pública de Navarra, 2 de la Universidad de Jaén y 6 sin especificar).

El tamaño de la muestra viene determinado por el tipo de muestreo, el parámetro a estimar, el error muestral admisible, la varianza poblacional y el nivel de confianza. La selección de los estudiantes se realizó por muestreo incidental (los sujetos de la URV). Los sujetos se identificaron automáticamente mediante un código proporcionado por el programa usado (Encuesta Fácil) y por su correo electrónico.

El test se realizó desde el mes de junio de 2013 hasta el mes de febrero de 2014 (la mayoría en noviembre) y el retest desde el mes de febrero hasta el mes de marzo de 2014 (la mayoría en febrero).

Según el análisis de resultados arrojado por Encuesta Fácil (2014) se entregaron 516 test, entendiendo como entregado el número de veces que los participantes pincharon sobre el enlace de la encuesta, y se contestaron 149 test, entendiendo por contestado el número de cuestionarios del test con al menos una pregunta contestada. Sin embargo, los sujetos que completaron el test hasta el final fueron 86.

Para animar a la participación en el test, se aprovechó la aplicación que tiene la página web de encuestas en línea para sorteos y se sorteó un lector electrónico entre todos los sujetos que terminaron el test y el retest. Con esta acción se pretendía motivar a los estudiantes a hacer el test de forma consciente y dar un valor añadido a la calidad de los resultados obtenidos.

El grupo de estudiantes que finalmente formaron la muestra (test) fue de 140 para el test ( $N_1=137$ ) y 56 para el retest ( $N_2=56$ ), elegidos de entre los estudiantes chinos matriculados en universidades españolas durante los cursos 2012-2013 y 2013-2014. La muestra del test ( $N_1=137$ ) presenta la siguiente distribución en función de las variables independientes asignadas:

1. Sexo: 23 hombres, 89 mujeres y 28 perdidos en el sistema.
2. Edad: con un rango de edad entre los 18 y los 40 años.
3. Nivel de español declarado: con un nivel de español diferenciado para este estudio que va del B1 al C2 (incluyendo los niveles oficiales en China EEE-4 y EEE-8, que equivaldrían a los niveles B2 y C1 del MCER).
4. Pareja española: la gran mayoría de los sujetos de la muestra no tienen ni han tenido pareja española. De los que sí han tenido, 6 son mujeres y 4 hombres.
5. Tiempo de residencia en España.

#### 4. Resultados y discusión

La investigación llevada a cabo, y expuesta a lo largo de estas páginas, ha dado como resultado final un "Test de nivel de comprensión pragmática para estudiantes universitarios sinohablantes de ELE", inexistente en la investigación en el ámbito de la enseñanza de ELE. Con este test ya es posible medir el nivel de comprensión pragmática que tienen los estudiantes de ELE. Se trata de un test estandarizado que cumple con los requisitos de utilidad (Bachman y Palmer, 1996) y estadísticos psicométricos de fiabilidad y validez. Las instrucciones del test se han redactado teniendo en cuenta que se trata de un test en línea y que las características técnicas de la web a través de la que se ha desarrollado tienen unas limitaciones. Una muestra del test final se reproduce a continuación:

#### Test de comprensión pragmática para estudiantes sinohablantes de español como lengua extranjera

##### Sección 1: Rutinas (uso de la lengua en situaciones concretas)

**Pregunta 1** Ana se toma un café en el bar y deja propina. El camarero le da las gracias.

*Camarero: Gracias.*

¿Qué crees que dirá Ana?

*Ana:*

- 1. «No, no.»
- 2. «Para qué.»
- 3. «Eso es todo.»
- 4. «De nada.»

**Pregunta 2** Ana ha ido a casa de Carlos a cenar con un grupo de amigos, ahora se están marchando y se despiden todos los amigos.

*Ana: Bueno, ha sido muy divertido. Gracias por la cena.*

¿Qué crees que dirá Carlos?

*Carlos:*

- 1. «Nada, nada. No ha sido nada.»
- 2. «Sí, estaba muy rica.»
- 3. «Otro día vamos al restaurante.»
- 4. «Yo también lo he pasado muy bien, a ver cuando repetimos.»

**Pregunta 3** Ana entra en el bar y ve a una compañera, Eva, llorando en una mesa.

Ana le pregunta qué le pasa.

*Eva: He suspendido el examen, ¡con lo que había estudiado!*

¿Qué crees que dirá Ana?

Ana:

- 1. «No llores, venga, puedes pedir una revisión de examen.»
- 2. «Bueno, vamos a comer helado toda la noche.»
- 3. «No sé, mala suerte.»
- 4. «Vamos a hablar con el profesor.»

## Sección II: Implicaturas (cuando hablamos de forma indirecta)

**Pregunta 1** Un compañero de clase de Ana, Miguel, está jugando a las cartas en el bar y Ana se acerca a saludarlo.

*Miguel: Hola, Ana. ¿Sabes jugar a las cartas?*

¿Qué crees que quiere decir Miguel?

- 1. Miguel quiere saber si Ana juega bien a las cartas.
- 2. Miguel quiere saber si a Ana le gusta jugar a las cartas.
- 3. Miguel quiere saber si Ana quiere jugar con ellos a las cartas.
- 4. Miguel quiere saber si Ana normalmente juega a las cartas.

**Pregunta 6** El Sr. López, el dueño del piso de Carlos, quiere subir el alquiler.

*Sr. López (dueño del piso): Carlos, a partir del mes que viene tengo que subiros el alquiler.*

*Carlos: ¡Hombre! El piso del año pasado era más grande y el alquiler era 100€ más barato.*

¿Qué crees que quiere decir Carlos?

- 1. Carlos cree que el piso es demasiado pequeño.
- 2. Carlos quiere volver a vivir al piso en el que vivía el año pasado.
- 3. Carlos ya no quiere vivir en ese piso con el alquiler más caro y se buscará otro piso.
- 4. Carlos intenta que el dueño del piso no le suba el alquiler.

**Pregunta 10** Ana le cuenta a su amigo Carlos que ha conocido a un chico que le gusta mucho en una fiesta.

Ana: *¿Tú lo conoces, Carlos?*

*Carlos: Sí, siempre anda contando historias raras.*

¿Qué piensa Carlos sobre el chico que le gusta a Ana?

- 1. Que es un chico con mucha imaginación.
- 2. Que es un chico simpático.

- 3. Que es un chico muy divertido.
- 4. Que no es un chico normal, y no lo ve muy apropiado para Ana.

**Pregunta 13** Ana estaba enferma y ha faltado a las últimas tres clases. Las clases son a las 8h de la mañana.

*Ana: Juan, no pude asistir a las tres últimas clases porque estuve enferma.*

*Juan (el profesor): Claro, claro, tuviste esa enfermedad que te deja pegada a la cama, ¿no? ¿Qué crees quiere decir el profesor?*

- 1. El profesor no estaba seguro, pero se imagina que Ana ha tenido esa enfermedad.
- 2. El profesor sabe que es una enfermedad grave y comprende que no haya podido ir a clase.
- 3. El profesor también ha tenido esa enfermedad y sabe bien lo que pasa.
- 4. El profesor cree que Ana está mintiendo.

### Sección III: Actos de habla (conversaciones cotidianas)

#### Instrucciones:

Lee la situación y decide qué opción te parece más natural, teniendo en cuenta lo que dice la otra persona:

1. Recuerda que tu respuesta debe tener sentido, es decir, debe ser parte coherente de lo que dice la otra persona.
2. Piensa que se trata de conversaciones normales entre personas normales (no son conversaciones divertidas, poco educadas, etc.).

**Pregunta 5** Ana va muy cargada y se encuentra con un compañero suyo de clase, Miguel, por el pasillo. Ana le pide ayuda a Miguel.

*Miguel: ¡Hola Ana! Vas muy cargada.*

¿Qué crees que probablemente dirá Ana?

Lee la respuesta de Miguel para estar seguro de que lo que Ana dice es lo adecuado.

*Ana:*

- 1. «Sí, ¿me echas una mano?»
- 3. «Sí, ya ves.»
- 3. «Como siempre.»
- 4. «Por supuesto.»

*Miguel: Claro, trae.*

**Pregunta 8** Eva necesita un bolígrafo para rellenar unos impresos y Carlos le presta el suyo. Es un bolígrafo muy especial, porque se lo regaló su madre cuando acabó el bachillerato. Eva sin darse cuenta se lleva el bolígrafo. Cuando Carlos y Eva se encuentran de nuevo, Carlos le pide el bolígrafo a Eva.

*Carlos: Eva, te quedaste con mi boli, ¿me lo devuelves?*

¿Qué crees que le dirá Eva?

*Eva:*

- 1. «Lo siento, lo he perdido. Te pagaré lo que cueste.»
- 2. «Lo siento mucho, lo he perdido. ¡Qué tonta soy!»
- 3. «Lo siento, lo he perdido. Bueno, solo era un boli, ¿no? Ya te daré el dinero, si quieres.»
- 4. «¡Ostras! Lo he perdido. Lo siento muchísimo. Ya te compraré otro, dime dónde lo compraste.»

Carlos: *Es que era un regalo de mi madre, no puedes comprarme otro.*

**Pregunta 11** Ana lleva ya unos meses en Tarragona y hace mucho que no ve a sus amigos y a su familia que están en China. Está hablando con una amiga y le dice lo mucho que le gustaría estar con ellos en ese momento.

Amiga: *¿Qué te pasa, Ana?*

¿Qué crees que dirá Ana a su amiga?

Lee la respuesta de su amiga para estar seguro de que lo que dice Ana es lo adecuado.

Ana:

- 1. «No sé, supongo que echo de menos a mi familia y a mis amigos.»
- 2. «No sé, supongo que necesito mucho a mi familia y a mis amigos de China.»
- 3. «No sé, es que tengo el corazón partido, entre China y España.»
- 4. «Es que mi familia y mis amigos de China están muy lejos.»

Amiga: *Es que están muy lejos, pero nos tienes a nosotros.*

## 4.1. Fiabilidad y validez del test: justificación estadística

### 4.1.1. Análisis de fiabilidad

Tal y como se ha visto anteriormente, la medida de la pragmática del español contempla tres subescalas aditivas que configuran la puntuación total: S1 (rutinas), S2 (implicaturas) y S3 (actos de habla). Cada una de ellas se compone de 18 ítems de distinta dificultad. La puntuación global teórica oscila entre 0 y 54 puntos. Sin embargo, en la muestra de baremación los límites obtenidos son:

#### FIGURA 2

Baremación de los límites obtenidos

MUESTRA DE BAREMACIÓN DE LOS LÍMITES OBTENIDOS						
ESCALA	N	MÍNIMO	MÁXIMO	MEDIA	DESV. TÍP.	ASIMETRÍA
S1rut	136	2	17	10,32	2,79	-,146
S2imp	135	2	17	9,45	3,84	-,011
S3ah	134	2	18	9,69	3,88	,162
Total	134	<b>9</b>	<b>49</b>	29,41	9,12	,051

En el total del test el mínimo de aciertos que se ha obtenido ha sido 9 sobre un total de 54, y el máximo, 49. Así, vemos que una persona que obtiene 9 o menos aciertos tiene un nivel de pragmática del español muy bajo, mientras que uno que obtiene 49 aciertos o más tiene un nivel de pragmática de español muy alto. La puntuación mínima es 2 en cada una de las escalas y, sin embargo, la mínima total es 9 y no 6 (2+2+2). Vemos, pues, que no todos los sujetos aciertan los mismos ítems “fáciles”, lo que incluye mayor variabilidad. El mismo razonamiento puede aplicarse a los ítems difíciles.

La fiabilidad como consistencia interna se refiere a la relación que tienen entre sí los distintos ítems que componen una escala al medir el mismo constructo, en este caso, la pragmática del español centrada en las rutinas, las implicaturas y los actos de habla respecto a la puntuación total. El coeficiente utilizado es el alfa de Cronbach. El valor obtenido en el alfa de Cronbach es de 0,85. Cuanto más se aproxime a su valor máximo, 1, mayor es la fiabilidad de la escala. Además, en determinados contextos y por tácito convenio, se considera que valores del alfa superiores a 0,7 o 0,8 (dependiendo de la fuente) son suficientes para garantizar la fiabilidad de la escala. La consistencia interna es el grado en el que los ítems se refieren o miden el mismo constructo. En función de la naturaleza del concepto a medir cabe esperar una mayor o menor consistencia interna.

En este caso, los valores obtenidos muestran que la consistencia interna de las escalas S2 (implicaturas) y S3 (actos de habla) es relativamente buena. Sin embargo, los valores de la S1 (rutinas) muestran que su consistencia interna resulta inequívocamente baja. Los valores son los que se muestran a continuación:

### FIGURA 3

Tabla del análisis de fiabilidad y consistencia interna de las subescalas

ESCALA	ALFA DE CROMBACH	Nº ÍTEMS	Nº SUJETOS
S1	0,56	18	136
S2	0,77	18	136
S3	0,77	18	134

Esto está en consonancia con el estudio Roever (2005) e Itomitsu (2009), que concluyen que el conocimiento de las rutinas (S1) va muy ligado a la exposición del aprendiente a la segunda lengua o a la lengua extranjera, mientras que en las otras secciones el vocabulario o el conocimiento gramatical pueden ayudar a descifrar la solución incluso por eliminación. Las rutinas no requieren solucionar problema alguno, sino más bien conocer de memoria la respuesta (Roever, 2005). El conocimiento de las implicaturas (S2) y de los actos de habla (S3), en cambio, está mucho más relacionado con el nivel de lengua.

Además, esto podría explicarse también por el modelo teórico, ya que tanto las implicaturas como los actos de habla (especialmente los perlocutivos) pueden tener un factor de comunicación indirecta que los hablantes tienen que descifrar y conllevan una variedad mucho más amplia de tareas, funciones, contenidos y contexto. Por el contrario, las rutinas son enunciados que los estudiantes repiten sin variación múltiples veces y les permiten relacionarse socialmente sin necesidad de una competencia lingüística elevada. La consistencia interna de este test es inferior a la obtenida por Roever; sin embargo, los estudios hechos por Roever (2005) e Itomitsu (2009) coinciden con estos resultados en cuanto a la fiabilidad obtenida, especialmente en las rutinas, que, igual que aquí, es sensiblemente menor que la obtenida en las secciones de implicaturas y actos de habla.

Se da el caso de que en la S3 (actos de habla) se ha detectado una pérdida de 2 sujetos, uno de los cuales se ha recuperado, pero la razón por la que no se ha podido recuperar el segundo es posiblemente por un fallo en el sistema, ya que el programa Encuesta Fácil puede que deje de registrar los resultados de aquellos participantes que no tengan activado el Java.

La fiabilidad como consistencia temporal se refiere a la estabilidad de las mediciones a lo largo del tiempo. En este caso, se obtuvo una submuestra ( $n=52$ ) que volvió a cumplimentar el test una vez transcurrido un periodo de tiempo aproximado de tres meses. En la siguiente tabla figuran las correspondientes correlaciones (coeficiente de correlación de Pearson) test-retest tanto para las subescalas como para la puntuación total.

#### FIGURA 4

Matriz de correlaciones test-retest

	S1 (TEST)	S2 (TEST)	S3 (TEST)	TOTAL (TEST)
S1 (retest)	0,590	0,471	0,590	0,574
S2 (retest)		0,737	0,623	0,770
S3 (retest)			0,489	0,565
Total (retest)				0,732

Como se puede observar, la correlación más alta corresponde a S2 (implicaturas), probablemente porque los estudiantes con una mayor competencia lingüística tienen también una mayor habilidad en reconocer las implicaturas. No obstante, la correlación más baja es la de los actos de habla, puesto que los actos de habla en un contexto en inmersión son la parte de la pragmática que más fácilmente se aprende con la interacción cotidiana y probablemente se da el caso de que los estudiantes los han aprendido durante los tres meses que han pasado entre el test y el retest. La correlación de la puntuación total tiene una fiabilidad test retest razonable ( $r_{xy}=0,73$ ).

#### 4.1.2. Análisis de validez

Para que el test sea válido hay que asegurarse de que está midiendo los parámetros que debe medir. Esto se demuestra con la selección de profesores expertos para llevar a cabo el papel de evaluador. Cada profesor evaluó una parte de la muestra. La evaluación a doble ciego mostró que la correlación entre las puntuaciones de pragmática (total) y la calificación de los jueces fue de 0,769 (rho de Spearman  $p < 0,01$ ). Por otro lado, el grado de acuerdo interevaluadores, calculado mediante el coeficiente de correlación ordinal tau b de Kendall, también resultó ser estadísticamente significativo ( $p < 0,01$ ) y con unos valores que oscilan entre 0,66 y 0,74. Esto indica que el acuerdo alcanzado difícilmente se puede atribuir al azar. La comprobación de las calificaciones otorgadas por los tres principales jueces evaluadores demuestra que los promedios de los tres son similares.

Los resultados de la replicación parcial de fiabilidad entre evaluadores garantizan en niveles muy satisfactorios la consistencia entre las réplicas. Las diferencias entre el test y el post-test evidencian que los resultados no varían de forma significativa, aunque el número de sujetos que respondieron el post-test fuera inferior.

Para reforzar la validez criterial se procede a comprobar que las calificaciones otorgadas por los principales evaluadores muestran diferencias estadísticamente significativas en las puntuaciones obtenidas por los sujetos en las subescalas y en la puntuación total final. Para ello, aplicamos un ANOVA de efectos aleatorios entre las calificaciones de los evaluadores, para ver hasta qué punto se hallan diferencias estadísticamente significativas. Entre paréntesis figura el tamaño de los grupos.

#### FIGURA 5

ANOVA entre las calificaciones obtenidas por los tres evaluadores en el test, para la variable puntuación total

	N	NIVEL BAJO	NIVEL MEDIO	NIVEL ALTO	F	SIG (P<)	ETA CUADRADO PARCIAL	POTENCIA
Evaluador 1	31	24,9 (11)	32,36 (11)	32,59 (10)	12,09	0,001	0,455	0,99
Evaluador 2	23	27,5 (4)	35,27 (11)	42 (8)	9,29	0,001	0,482	0,95
Evaluador 3	28	23,5 (12)	34,18 (11)	40,2 (5)	11,8	0,001	0,486	0,98

Como se puede apreciar, a pesar de contar con tamaños muestrales diferentes en cada uno de los evaluadores y para cada una de las tres calificaciones (bajo, medio y alto), los promedios en cada uno de ellos son similares, así como los valores F (estadístico de contraste) y el tamaño del efecto (operativizado por eta cuadrado parcial), la capacidad discriminadora de los

evaluadores de acuerdo con las puntuaciones obtenidas en el test y la potencia de la prueba (la capacidad de detectar diferencias entre los promedios).

A continuación, se muestra el grado de acuerdo entre los evaluadores en función de su poder discriminativo en cada una de las subescalas.

### FIGURA 6

ANOVA sobre el grado de acuerdo entre evaluadores Subescala 1

	N	NIVEL BAJO	NIVEL MEDIO	NIVEL ALTO	F	SIG (P<)	ETA CUADRADO PARCIAL	POTENCIA
Evaluador 1	31	9,45 (11)	11,09 (11)	12,07 (10)	3,25	0,053	0,184	0,574
Evaluador 2	23	8,5 (4)	11,72 (11)	13 (8)	2,96	0,075	0,229	0,512
Evaluador 3	28	9 (12)	10,8 (11)	13 (5)	3,53	0,045	0,22	0,60

Tal y como se puede observar, las tres calificaciones de dos de los tres evaluadores no se distinguen en los promedios obtenidos en S1. Los estadísticos obtenidos son similares y ambos están muy próximos al límite considerado para obtener la significación estadística (0,05). Si observamos el tamaño del efecto obtenido en todos ellos, así como la potencia estadística, es fácil concluir que un aumento de los sujetos evaluados facilitarían alcanzar la significación estadística.

### FIGURA 7

ANOVA sobre el grado de acuerdo entre evaluadores Subescala 2

	N	NIVEL BAJO	NIVEL MEDIO	NIVEL ALTO	F	SIG (P<)	ETA CUADRADO PARCIAL	POTENCIA
Evaluador 1	31	8 (11)	9,91 (11)	13,70 (10)	9,22	0,001	0,389	0,96
Evaluador 2	23	10,75 (4)	11,27 (11)	13,75 (8)	3,137	0,065	0,237	0,539
Evaluador 3	28	8,45 (12)	11,83 (11)	14,5 (5)	9,93	0,001	0,443	0,97

En este caso, dos de los tres evaluadores obtienen diferencias estadísticamente significativas entre los promedios de sus evaluados. Cabe destacar las similitudes encontradas entre los evaluadores 2 en todos los estadísticos calculados. De nuevo, de manera similar al caso anterior, el número bajo de sujetos evaluados dificulta el hallazgo de diferencias estadísticamente significativas.

**FIGURA 8**

ANOVA sobre el grado de acuerdo entre evaluadores Subescala 3

	N	NIVEL BAJO	NIVEL MEDIO	NIVEL ALTO	F	SIG (P<)	ETA CUADRADO PARCIAL	POTENCIA
Evaluador 1	31	7,45 (11)	11,36 (11)	14,9 (10)	13,61	0,001	0,48	0,99
Evaluador 2	23	8,5 (4)	12,27 (11)	15,25 (8)	11,67	0,001	0,539	0,98
Evaluador 3	28	7,08 (12)	12,09 (11)	13,04 (5)	8,32	0,002	0,44	0,94

En la última subescala, la S<sub>3</sub> (actos de habla), volvemos a encontrar acuerdo unánime entre los tres evaluadores, a nivel de significación estadística, que se aleja sensiblemente del valor mínimo exigible (0,05). El análisis del acuerdo entre evaluadores en función de su poder discriminativo para cada una de las escalas nos indica que la S<sub>3</sub> (actos de habla) contribuye a generar una apreciación unánime entre ellos, al discriminar los tres con niveles de significación estadística exigente y tamaños de efecto grandes. En segundo lugar, tenemos la subescala S<sub>2</sub> (implicaturas), donde el acuerdo del poder discriminante se mantiene en dos de los tres evaluadores, también con un nivel de exigencia similar. Por último, la subescala S<sub>1</sub> parece *robusta* frente a la discriminación por parte de los evaluadores: solo uno de ellos consigue diferenciar los promedios respecto a sus calificados como bajo, medio o alto. No obstante, como ya se ha comentado anteriormente, con un número mayor de sujetos evaluados, se aumentaría el poder estadístico de la prueba con el consiguiente hallazgo de diferencias estadísticamente significativas.

Desde el modelo teórico, se repiten los factores de discriminación que se han dado en los resultados de fiabilidad como consistencia interna, de forma lógica por los mismos motivos. Parece ser que los actos de habla (S<sub>3</sub>) son los elementos visibles de uso de la pragmática más fáciles de detectar, ya que son los que están más relacionados con el nivel de dominio lingüístico, seguidos de las implicaturas (S<sub>2</sub>). Tanto los actos de habla como las implicaturas se asocian en muchos casos a un mejor dominio de la lengua, y los errores cometidos con el uso de los mismos son más evidentes.

Finalmente, como apoyo adicional a la validez de constructo de la escala elaborada para medir el nivel de pragmática en ELE, presentamos otros cálculos correlacionales que deberían mostrar valores positivos y moderados, siempre diferentes y estadísticamente significativos de 0. Estos cálculos son los siguientes:

- (a) Sexo: Aunque existen pocos estudios sobre el tema y la mayoría de ellos no profundizan en él, la poca bibliografía encontrada indica que se dan diferencias entre hombres y mujeres en el aprendizaje de la pragmática de la interlengua (Herbert, 1990; Yu, 2004). Los cálculos realizados al efecto mediante ANOVA, de efectos fijos, van en esta dirección:  $F_{1,109} = 4,67$ ;

$p < 0,034$ . Las mujeres tienen un dominio ligeramente mayor de pragmática que los hombres. No obstante, el tamaño del efecto nos indica que el sexo solo explica un 4,2% de la varianza de las puntuaciones en el test total y sería pertinente realizar un estudio más extenso sobre el papel que tiene el sexo de los aprendices en la adquisición de la competencia pragmática.

### FIGURA 9

Estadísticos descriptivos. Puntuación total

SEXO	MEDIA	DESV. TÍP.	N
Hombre	27,2857	8,71288	21
Mujer	31,8409	8,67138	88
Total	30,9633	8,82539	109

- (b) Nivel declarado de español: Aunque el nivel de español declarado no coincide totalmente con el dominio de la pragmática, cabría esperar una correlación positiva y moderada entre ambas variables: a mayor nivel de español declarado, mayor nivel de pragmática. Sin embargo, esta asociación no debe ser especialmente intensa, ya que, si lo fuese, este trabajo no tendría sentido, no habría necesidad de medir la pragmática por separado. Los cálculos realizados apoyan esta predicción:  $r_{xy} = .0,329$  ( $n=109$ );  $p < 0,001$ . Un 10,8% ( $0,329^2 = 0,108$ ) de la varianza de las puntuaciones de la pragmática se explican por los diferentes niveles de español declarado.
- (c) Tiempo de residencia en España: Es fácil entender que aquellos que lleven más tiempo en España habrán mejorado su pragmática por encima de aquellos que lleven menos tiempo en el país. Los cálculos realizados (coeficiente de correlación de Pearson) apoyan esta predicción:  $r_{xy} = .0,458$  ( $n=109$ );  $p < 0,001$ . Esto explica  $0,458^2 = 0,209$  (20,9%). Esto es, un 20,9% de la varianza de las puntuaciones de la pragmática se explican por el tiempo de residencia en España.
- (d) Edad: Esta variable biológica no debería mostrar ningún tipo de correlación con la pragmática, a menos que esté asociada a otras variables anteriores (como nivel de castellano declarado o tiempo de permanencia en España), actuando estas como variables confundentes o “contaminadoras”. Para ello, se neutralizan estas mediante el cálculo del coeficiente de correlación parcial. El resultado confirma la predicción expuesta:  $r_{xy} = .0,093$  ( $g.l.=105$ );  $p > 0,33$ . Esto explica  $0,093^2 = 0,008$  (menos del 1%).

## 5. Conclusiones

Tanto los resultados estadísticos obtenidos como el análisis de las cualidades que definen la utilidad de un test de lengua evidencian que es posible crear un test útil, fiable y validado que mida el nivel de comprensión pragmlingüística de los estudiantes sinohablantes de ELE:

(1) El análisis de las cualidades del test que evalúan su utilidad muestra que el resultado de la investigación es un test útil que cumple todas las cualidades necesarias para que lo sea: autenticidad, validez del constructo, interactividad, impacto y practicabilidad. Todas las cualidades cumplen con los requisitos necesarios de forma efectiva.

(2) El análisis de fiabilidad del test muestra que este está midiendo, en efecto, el constructo definido en las tres subescalas propuestas: las rutinas, las implicaturas y los actos de habla. El coeficiente del alfa de Cronbach (0,85) implica que los ítems que componen cada sección obtienen un resultado aceptable de fiabilidad como consistencia interna.

(3) En cuanto a la validez del test, las pruebas realizadas que se hicieron evidencian que existe una correlación en la puntuación obtenida en el test y la otorgada por jueces expertos a una parte heterogénea de la muestra (n=34).

## 6. Bibliografía citada

AUSTIN, John Langshaw, 1976 [1962]: *How to do things with words: The william james lectures delivered at harvard university in 1955*, Oxford: Oxford University Press.

BACHMAN, Lyle F., 1990: *Fundamental considerations in language testing*, Oxford: Oxford University Press.

BACHMAN, Lyle F., y Adrian S. PALMER, 1996: *Language testing in practice: Designing and developing useful language tests*, Oxford: Oxford University Press.

BACHMAN, Lyle F., y Adrian S. PALMER, 2010: *Language assessment in practice*, Oxford: Oxford University Press.

BARDOVI-HARLIG, Kathleen, y Zoltán DÖRNYEI, 1998: "Do language learners recognize pragmatic violations? Pragmatic versus grammatical awareness in instructed L2 learning", *TESOL Quarterly* 32 (2), 233-259.

BROWN, James D., 2001: "Pragmatic tests: Different purposes, different tests" en Kenneth R. ROSE y Gabriele KASPER (eds.): *Pragmatics in language teaching*, Cambridge: Cambridge University Press, 301-326.

BROWN, James D., y Thom HUDSON, 2002: *Criterion-referenced language testing*, Cambridge: Cambridge University Press.

CANALE, Michael, 1983: "On some dimensions of language proficiency" en John OLLER (ed.): *Issues in language testing research*, Rowley, Massachusetts: Newbury House.

CANALE, Michael, y Merrill SWAIN, 1980: "Theoretical bases of communicative approaches to second language teaching and testing", *Applied Linguistics* 1 (1), 1-47.

FARHADY, Hossein, 1980: *Justification, development and validation of functional language tests* (unpublished). University of California, Los Angeles.

GUMPERZ, John J., y Dell HYMES (eds.) 1972: *Directions in sociolinguistics*, New York: Holt.

HERBERT, Robert K., 1990: "Sex-based differences in compliment behaviour", *Language in Society* 19, 201-224.

HUDSON, Thom, Emily DETMER y James BROWN, 1992: *A framework for testing cross-cultural pragmatics*, Honolulu: Second Language Teaching Curriculum Center, University of Hawai'i.

HUDSON, Thom, Emily DETMER y James BROWN, 1995: *Developing prototypic measures of cross-cultural pragmatics*, Honolulu: Second Language Teaching y Curriculum Center, University of Hawai'i at Manoa.

INSTITUTO CERVANTES, 2007: *Plan curricular del instituto cervantes: Niveles de referencia para el español* (2.ª ed.), Madrid: Biblioteca Nueva.

ITOMITSU, Masayuki, 2009: *Developing a test of pragmatics of Japanese as a foreign language*. Unpublished Ph.D. Dissertation, Ohio State University [<https://www.semanticscholar.org/paper/Developing-a-Test-of-Pragmatics-of-Japanese-as-a-Itomitsu/c7d86f5af7065a7798f786ebcb884f6554171bf8>].

JOHNSTON, Bill, Gabrielle KASPER y Steven ROSS, 1998: "Effect of rejoinders in production questionnaires", *Applied Linguistics* 19 (2), 157-182.

KASPER, Gabrielle (ed.), 1995: *Pragmatics of Chinese as native and target language*, Honolulu: University of Hawai'i at Manoa.

McNAMARA, Tim, y Carsten ROEVER, 2006: *Language testing: The social dimension*, Malden: Blackwell.

NORTH, Brian, 2000: *The Development of a Common Framework Scale Proficiency*, Frankfurt: Peter Lang.

OLLER, John, 1979: *Language tests at school: A pragmatic approach* (4.ª ed.), Harlow: Longman.

ROEVER, Carsten, 2000: "Web-based language testing: opportunities and challenges", *22nd Annual Language Testing Research Colloquium*, Vancouver, BC, Canada.

ROEVER, Carsten, 2001: *A web-based test of interlanguage pragmalinguistic knowledge: Speech acts, routines, and implicatures*. Unpublished Ph.D. Dissertation, University of Hawai'i [[https://books.google.es/books/about/A\\_Web\\_based\\_Test\\_of\\_Interlanguage\\_Pragma.html?id=FFsiOAAACAAJ&redir\\_esc=y](https://books.google.es/books/about/A_Web_based_Test_of_Interlanguage_Pragma.html?id=FFsiOAAACAAJ&redir_esc=y)].

ROEVER, Carsten, 2005: *Testing ESL pragmatics*, Frankfurt am Main: Peter Lang.

SCHAUER, Gila, A., 2009: *Interlanguage pragmatic development: the study abroad context*, London: Continuum [http://eprints.lancs.ac.uk/27262/].

SHOHAMY, Elana, 1996: "Competence and performance in language testing" en Gillian BROWN, Kirsten MALMKJÆR y John WILLIAMS (eds.): *Performance and competence in second language acquisition*, Cambridge: Cambridge University Press, 136-151.

TAGUCHI, Naoko, 2008: "Cognition, language contact, and the development of pragmatic comprehension in a study-abroad context", *Language Learning* 58 (1), 33-71.

THORNDIKE, Robert, y Elizabeth HAGEN, 1980: *Test y técnicas de medición en psicología y educación: Elaboración, diseños, investigación, aplicación*, Mexico: Trillas.

YAMASHITA, Sayoko, 2008: "Investigating interlanguage pragmatic ability: What are we testing?" en Eva ALCÓN y Alicia MARTÍNEZ (eds.): *Investigating pragmatics in foreign language learning, teaching and testing*, Bristol: Multilingual Matters, 201-223

Yu, Ming-Chung, 2004: "Interlinguistic variation and similarity in second language speech act behavior", *The Modern Language Journal* 88 (i), 102-117.

## 7. Anexos

Ilustración 1: Ejemplo del tipo de tarea de la Sección 1 (Rutinas)

Sección 1: *Rutinas conversacionales (cómo hablamos en situaciones concretas)*, se describe una situación y, a continuación, en un diálogo en el que se conoce una réplica y la otra está incompleta el candidato debe elegir la respuesta adecuada para completar la réplica:

**TEST DE ESPAÑOL COMO LENGUA EXTRANJERA PARA ESTUDIANTES CHINOS: GRAMÁTICA Y PRAGMÁTICA**

Abandonar-> Continuaré más tarde

Clica en la opción correcta y luego en el botón "SIGUIENTE".

-Puedes parar cuando quieras y continuar en otro momento dándole al botón "CONTINUARÉ MÁS TARDE".

RECUERDA: Si le das al botón "ABANDONAR" ya no podrás seguir.

IMPORTANTE:  
- Por favor, si tienes alguna duda NO PREGUNTES A NADIE, para el estudio es imprescindible que lo hagas solo.

**\*PREGUNTA 1**

Ana se toma un café en el bar y deja propina. El camarero le da las gracias.

**Camarero: Gracias.**

¿Qué crees que dirá Ana?

Ana:

1. "No, no."  
 2. "Para qué."  
 3. "Eso es todo."  
 4. "De nada."

<-Anterior    Siguiente->

Ilustración 2: Ejemplo del tipo de tarea de la Sección 2 (Implicaturas)

Sección 2: *Implicaturas (cuando hablamos de forma indirecta)*, se describe una situación para contextualizar y, a continuación, se presenta un diálogo completo de dos réplicas. El candidato tiene que interpretar qué quiere decir una de las personas que interactúan en el diálogo.

**TEST DE ESPAÑOL COMO LENGUA EXTRANJERA PARA ESTUDIANTES CHINOS: GRAMÁTICA Y PRAGMÁTICA**

Abandonar-> Continuaré más tarde

Clica en la opción correcta y luego en el botón "siguiente".

-Puedes parar cuando quieras y continuar en otro momento dándole al botón "Continuaré más tarde".

RECUERDA: Si le das al botón "Abandonar" ya no podrás seguir.

IMPORTANTE:  
- Por favor, si tienes alguna duda NO PREGUNTES A NADIE, para el estudio es imprescindible que lo hagas solo.

**\*PREGUNTA 1**

Ana está comiendo en un restaurante.

Ana (a la mesa del lado): Perdonen, ¿necesitan la sal?

¿Qué crees que quiere decir Ana?

1. Quiere saber si a la comida de los demás clientes también les falta sal.

2. Quiere saber si los comensales de la otra mesa ya se han puesto sal.

3. Quiere saber si puede usar el salero de la otra mesa.

4. Quiere saber si los comensales de la otra mesa comen con sal.

<-Anterior Siguiente->

71%

Ilustración 3: Ejemplo del tipo de tarea de la Sección 3 (Actos de habla)

Sección 3: *Actos de habla (conversaciones cotidianas)*, se describe una situación para contextualizar y, a continuación, se presenta un diálogo de dos réplicas con una sola réplica. El candidato debe elegir la réplica que considera más adecuada para esa situación.

**TEST DE ESPAÑOL COMO LENGUA EXTRANJERA PARA ESTUDIANTES CHINOS: GRAMÁTICA Y PRAGMÁTICA**

Abandonar-> Continuaré más tarde

- Lee la situación y decide qué opción te parece más natural, teniendo en cuenta lo que dice la otra persona:

1. Recuerda que tu respuesta debe tener sentido, es decir, debe ser parte coherente de lo que dice la otra persona.

2. Piensa que se trata de conversaciones normales entre personas normales (no son conversaciones divertidas, poco educadas, etc.).

Clica en la opción correcta y luego en el botón "siguiente".

-Puedes parar cuando quieras y continuar en otro momento dándole al botón "Continuaré más tarde".

RECUERDA: Si le das al botón "Abandonar" ya no podrás seguir.

**\*PREGUNTA 2**

Carlos llega a casa y se encuentra a su vecina, una señora mayor, que va cargada con la compra. Se ofrece a ayudarla.

¿Qué crees que probablemente dirá Carlos?

Lee la respuesta de la vecina para estar seguro de que lo que Carlos dice es lo adecuado.

Carlos:

1. "Disculpe, ¿podría ayudarme a llevar la compra?"

2. "Déjeme que la ayude."

3. "¿Le importa que le suba la compra? Las bolsas pesan mucho y yo soy fuerte."

4. "¿Te ayudo?"

Vecina: "Gracias, chico..."

<-Anterior Siguiente->

87%