La selección léxica en pretérito imperfecto y pretérito indefinido en un corpus escrito de alumnos japoneses de ELE

Carmen Muñoz Sanz
Universidad Complutense de Madrid

En el presente trabajo partiremos del análisis de un corpus escrito por alumnos japoneses de español para examinar la selección léxica en pretérito imperfecto e indefinido así como los errores encontrados. La estructura a seguir será la siguiente. En primer lugar, trataremos de forma muy concisa las diferencias entre los dos tiempos verbales que nos ocupan. Seguidamente presentaremos las características del corpus con el que hemos trabajado. A continuación, nos detendremos en explicar el análisis de errores realizado, observando especialmente la tipología de errores seguida. Posteriormente expondremos y trataremos los datos extraídos a partir del análisis del corpus relativos a la selección léxica. Por último, finalizaremos ofreciendo las conclusiones de nuestro trabajo.

Comencemos, pues, tratando los dos tiempos verbales que nos ocupan: el pretérito indefinido y el pretérito imperfecto. No resulta sorprendente comprobar que, en general, diferenciar el uso de estos tiempos supone una dificultad para los alumnos japoneses de español, como veremos más adelante. ¿Cómo podemos explicar la diferencia entre ambos? Se trata de una cuestión tan compleja que no encontramos un acuerdo entre los propios gramáticos. Fundamentalmente podemos considerar que existen dos posturas enfrentadas. Por un lado, aquellos autores que consideran que se trata de una diferencia temporal, como por ejemplo Bello (1984), Guitérrez Araus (1995), Porto Dapena (1989), Rojo (1990) y Veiga (2008), y, por otro, los que consideran que la diferencia es aspectual, destacando Carrasco Gutiérrez (1999), García Fernández (2004), Granados *et al*. (1979), Moreno de Alba (2006), entre muchos otros. Sin menospreciar la otra propuesta, personalmente nos mostramos de acuerdo con el segundo grupo de gramáticos y consideramos la diferencia entre estos tiempos verbales como aspectual. De este modo, el pretérito indefinido expresaría aspecto perfectivo mientras que el imperfecto lo expresaría imperfectivo. Sin embargo, por su complejidad, esta diferencia no sería suficiente para explicar a los alumnos de ELE cuándo se usa cada tiempo verbal, de manera que los profesores debemos recurrir en muchas ocasiones a la simplificación y a tratar los usos de cada tiempo de forma independiente.

A continuación trataremos brevemente las características del corpus con el que hemos trabajado. Dicho corpus fue recopilado durante una estancia de investigación en la Universidad de Estudios Extranjeros de Tokio (TUFS) y no habría sido posible este trabajo sin la amable colaboración de varios profesores de español[1]. Se trata de un corpus que recoge el habla escrita de alumnos universitarios japoneses de español. Concretamente, hemos contado con la participación de 148 informantes pertenecientes a tres universidades japonesas (la Universidad de Estudios Extranjeros de Tokio, la

1. Deseamos expresar nuestro agradecimiento a los profesores Toshihiro Takagaki, Shigenobu Kawakami, Kimiyo Nishimura, Concha Moreno y Arturo Varón.

Universidad de Sophia y la Universidad de Kanagawa). Los informantes se han dividido en cuatro grupos atendiendo al tiempo que llevaban estudiando español, formando el grupo G1 aquellos que habían estudiado español durante un año, el G2 los estudiantes que han estudiado durante dos años, y así sucesivamente hasta el grupo G4, formado por estudiantes que han estado aprendiendo español durante cuatro o más años.

Para la recopilación del corpus, se pidió a los informantes que escribieran una redacción de aproximadamente diez líneas de extensión sobre uno de los tres temas propuestos. Estos tres temas se seleccionaron con la intención de ser motivadores para los informantes y ofrecerles la posibilidad de expresarse en pasado. Los temas elegidos son los siguientes: *un día divertido que pasaste con tus amigos*, *un viaje o unas vacaciones interesantes* y *habla sobre las clases de español del curso pasado*. Obviamente, la elección de un tema u otro condiciona la selección léxica, ya que verbos como *estudiar* o *aprender* se asocian con el tercer tema y no con los demás.

La extensión total del corpus recogido es de 13.020 palabras. La prueba se presentó como una colaboración voluntaria y los informantes la realizaron sin supervisión. El plazo aproximado para su compleción fue de dos semanas. Por supuesto, la transcripción del corpus se realizó sin alterar los originales en ningún caso, de manera que se han mantenido todas las erratas o errores que pudieran aparecer en las redacciones de los informantes. Cada uno fue identificado con un código que indica datos personales del informante como el sexo, la edad o el tiempo de estudio de español, pero manteniendo siempre el anonimato. El corpus completo puede consultarse en Muñoz Sanz (2017). Las redacciones fueron anotadas de forma manual con símbolos que marcaban todo aquello que se deseaba analizar posteriormente, como cada uno de los tiempos del pasado. Asimismo, se señalaba si en los ejemplos aparecían errores o por el contrario eran correctos.

Dado que la manera de analizar y clasificar los errores puede variar ampliamente dependiendo de la metodología seguida, nos detendremos brevemente en señalar los tipos de errores a los que se ha prestado atención en nuestro trabajo. Una vez que se identificaban los ejemplos en algún tiempo verbal de pasado, se examinaba si estos eran correctos en el uso o no, es decir, su adecuación al uso. Posteriormente se analizaba la existencia de posibles errores adicionales, como errores en la conjugación y equivocaciones en la persona. Por ejemplo, en la oración *yo viajó por Shizuoka*, el pretérito indefinido es correcto –de manera que el ejemplo se marcaría como no erróneo– pero aparece una confusión en la persona –*viajó* por *viajé*–, por lo que se indicaría que existe un error de persona. Sin embargo, no se han tenido en cuenta errores ortográficos ni de otro tipo (confusión entre *l* y *r*, tildes, etc.).

Una vez tratada la tipología de errores seguida, comenzaremos la exposición de los datos extraídos del corpus. En primer lugar, observemos el recuento de ejemplos de verbos de pasado:

Total de verbos en pasado	1614
Verbos en pretérito indefinido	1080
Verbos en pretérito imperfecto	425
Verbos en pretérito perfecto	51
Verbos en pretérito pluscuamperfecto	55
Verbos en pretérito anterior	3

Tabla 1. Número total de casos de los verbos en el corpus

En la tabla 1 se aprecia que el tiempo verbal de pasado más frecuente en nuestro corpus es el pretérito indefinido, con 1080 casos, lo que supone un 67% del total de verbos en pasado, muy alejado del 26% del imperfecto, con 425 casos. De este total de apariciones de cada tiempo verbal, ¿cuántas corresponden a verbos distintos? Es decir, ¿cuántos verbos diferentes –*ser*, *hacer*, *ir*, etc.– aparecen en nuestro corpus?

Podemos observar el número de verbos distintos que aparecen en pretérito indefinido en la tabla 2, que asimismo recoge los casos totales divididos en grupos y el promedio resultante:

	G1	G2	G3	G4
Casos totales	477	314	119	170
Verbos distintos	87	76	61	68
Promedio	5,5	4,1	1,9	2,5

Tabla 2. Número de verbos distintos en pretérito indefinido

Como se puede apreciar, el grupo G3 es el que presenta mayor variedad léxica, puesto que los informantes han usado más verbos diferentes proporcionalmente. Le sigue en variedad el grupo G4, lo que indica que en general los informantes con mayor nivel de español son capaces de utilizar más verbos distintos. La tabla 3 recoge los mismos datos pero relativos al pretérito imperfecto:

	G1	G2	G3	G4
Casos totales	151	126	82	66
Verbos distintos	25	29	27	21
Promedio	6	4,3	3	3,1

Tabla 3. Número de verbos distintos en pretérito imperfecto

Nuevamente, la variedad léxica aumenta a medida que aumenta el nivel de los informantes, con los grupos G3 y G4 obteniendo resultados prácticamente idénticos.

Tras haber mostrado los datos totales y el promedio de verbos diferentes, es necesario examinar cuántos errores de uso encontramos en el corpus. En primer lugar, ofrecemos una tabla que recoge la evolución en los porcentajes de aciertos en los dos tiempos verbales:

Grupo	Indefinido	Imperfecto
G1	79,25%	80,13%
G2	85,67%	78,57%
G3	70,59%	75,61%
G4	87,06%	90,91%

Tabla 4. Porcentajes de aciertos totales en cada tiempo verbal

Nótese que tan solo aparecen en la tabla 4 los casos que son correctos en el uso y que además no presentan errores adicionales de conjugación. En general, los porcentajes de aciertos son bastante altos, pero no debemos olvidar que se trata de un corpus escrito y que los informantes han tenido tiempo suficiente para escribir sus redacciones, de manera que pueden haber seleccionado ejemplos de los que estén seguros y evitado casos dudosos. Sin embargo, no sería adecuado considerar que el indefinido y el imperfecto no presentan problemas para los informantes. Estos dos tiempos verbales se confunden entre sí en un porcentaje elevadísimo, como muestra la tabla 5:

Grupo	Indefinido	Imperfecto
G1	87,76%	85,19%
G2	100%	75%
G3	88,24%	94,12%
G4	83,34%	83,34%

Tabla 5. Porcentaje de confusión entre el pretérito indefinido y el imperfecto

Por ejemplo, de los errores de uso del pretérito indefinido en el grupo G1, un 87,76% corresponden a una confusión con el pretérito imperfecto. Es decir, el 87,76% de los errores habrían sido correctos si los informantes hubieran utilizado el imperfecto en lugar del indefinido. Los elevado de estos porcentajes en todos los grupos pone en evidencia los problemas que tienen los informantes para distinguir entre indefinido e imperfecto.

Seguidamente mostramos la tabla que recoge los diez verbos más frecuentes en pretérito indefinido. Los datos se desglosan por grupos y también se ofrece el total. Aquellos verbos que son correctos en el uso se marcan en la columna O, mientras que los incorrectos aparecen en la columna X.

	G1		**G2**		**G3**		**G4**		**Total**	
	O	X	O	X	O	X	O	X	O	X
Ir	94	-	70	-	19	-	26	-	209	-
Ser	40	18	26	9	5	2	14	7	85	36
Comer	38	1	14	-	2	-	5	-	59	1
Gustar	14	-	10	-	1	-	6	-	31	-
Poder	16	3	4	1	1	1	4	2	25	7
Pasar	9	-	8	1	10	-	2	-	29	1
Ver	15	1	9	1	2	-	1	1	27	3
Hablar	12	2	10	-	2	-	3	-	27	2
Visitar	11	-	7	-	2	-	4	-	24	-
Comprar	14	-	5	-	2	-	2	-	23	-

Tabla 6. Diez verbos más frecuentes en pretérito indefinido

Como se puede apreciar, el verbo más frecuente es *ir*, con un total de 209 casos. Resulta llamativo que no aparezca ni un solo error en este verbo en ninguno de los grupos. El siguiente verbo según su frecuencia de aparición es *ser*, con 85 casos correctos y 36 incorrectos.

La tabla 7 recoge los diez verbos más frecuentes en pretérito imperfecto. Al igual que la tabla 6, los casos correctos se indican con O y los incorrectos con X.

	G1		**G2**		**G3**		**G4**		**Total**	
	O	X	O	X	O	X	O	X	O	X
Ser	30	3	28	4	16	2	15	1	89	10
Estar	23	6	20	1	11	2	7	-	61	9
Haber	22	-	13	-	7	-	4	-	46	-
Tener	11	2	10	-	10	2	6	1	37	5

Hacer	10	-	5	-	-	-	4	-	19	-
Querer	7	-	5	-	-	-	3	-	15	-
Gustar	4	1	3	1	2	-	1	2	10	4
Poder	-	-	5	1	1	-	2	1	8	2
Parecer	2	1	2	-	1	-	-	1	5	2
Pensar	-	-	1	1	3	-	1	-	5	1

Tabla 7. Diez verbos más frecuentes en pretérito imperfecto

En este caso, el verbo *ir* –tan frecuente en pretérito indefinido– ni siquiera aparece en la tabla. El verbo más frecuente en imperfecto es *ser*, con 89 casos correctos y 10 incorrectos. Va seguido de *estar*, que obtiene 61 casos correctos y 9 incorrectos.

Para que sea más fácil de visualizar, podemos reflejar estos datos de la siguiente manera: calculando el porcentaje de errores de uso para cada verbo en cada uno de los tiempos verbales. Las tablas 8 y 9 muestran dichos porcentajes:

Indefinido	
Ir	0%
Ser	42,3%
Comer	1,7%
Gustar	0%
Poder	28%
Pasar	3,4%
Ver	11,1%
Hablar	7,4%
Visitar	0%
Comprar	0%

Imperfecto	
Ser	11,2%
Estar	14,7%
Haber	0%
Tener	13,5%
Hacer	0%
Querer	0%
Gustar	40%
Poder	25%
Parecer	40%
Pensar	20%

Tablas 8 y 9. Porcentajes de errores de uso en pretérito indefinido e imperfecto

A pesar de los elevados porcentajes obtenidos por verbos como *gustar* o *parecer* en pretérito imperfecto, consideramos que no son tan relevantes ya que el número total de casos en estos verbos no es significativo. Sin embargo, nos parece especialmente llamativo el caso del verbo *ser* en ambos tiempos verbales, así como de *estar* y *tener* en pretérito imperfecto.

Ya que los verbos *ser*, *estar* y *haber* en pretérito imperfecto son los más frecuentes en nuestro corpus, hemos observado qué porcentaje suponen del total de verbos en cada grupo de informantes. Dichos porcentajes se recogen en la tabla que aparece a continuación:

G1	G2	G3	G4
49,7%	48,4%	41,5%	39,4%

Tabla 10. Porcentaje total de *ser, estar* y *haber* en pretérito imperfecto

Nótese que estos verbos suelen ir asociados con la descripción. Pues bien, como se puede apreciar en la tabla 10, a medida que avanza el nivel de los informantes, el porcentaje conjunto de estos tres verbos se reduce. Esto supone que los alumnos de mayor nivel son capaces de utilizar más verbos diferentes en pretérito imperfecto. Si atendemos a los errores que aparecen en este tiempo verbal, como ya se trató en la tabla 4, también se puede comprobar que no solo aumenta la variedad léxica sino también se reducen los errores de uso.

Por otra parte, se ha apreciado que el pretérito indefinido presenta un mayor número de errores de conjugación que el imperfecto. Este dato no debería sorprendernos si tenemos en cuenta que este tiempo verbal tiene una conjugación con muchas irregularidades, especialmente si se compara con la del imperfecto, que es relativamente simple. Sin embargo, al examinar qué verbos concretos presentaban dichos errores, los resultados no han sido los esperados. Los siete verbos en pretérito indefinido con más errores de conjugación son los siguientes: *comer* y *andar* (7 casos cada uno), *poder* (6 casos), *hablar* (4 casos) y *hacer*, *disfrutar* y *llegar* con 2 casos. Después encontramos verbos con un caso cada uno, como *comprar*, *pasar* o *beber*.

Resulta llamativo que en la lista aparezcan tantos verbos regulares, que en principio no deberían dar problemas a los informantes. Un análisis detallado de los datos ha esclarecido el motivo de esta aparente contradicción. Un elevadísimo porcentaje de los errores de conjugación se produce en la primera persona del plural (*nosotros*), como se expresa en la tabla 11:

	G1	G2	G3	G4
Total errores de conjugación	26	15	8	3
Errores en 1ª p. pl.	19	12	5	2
Porcentaje 1ªp. pl. sobre total	67,8%	80%	62,5%	66,7%

Tabla 11. Número y porcentaje de errores de conjugación en pretérito indefinido.

Como se puede apreciar, estos porcentajes son lo suficientemente significativos como para dotar de importancia a este fenómeno. Es decir, parece que los informantes de nuestro corpus tienen más dificultades para conjugar correctamente la persona *nosotros* que el resto de personas. Esto puede estar provocado por la conjugación de los verbos regulares de la primera y la tercera conjugación, que es igual al presente en esta persona (*ayer cantamos/ hoy cantamos*), lo que podría conducir a los informantes a cometer equivocaciones.

Una vez observados los datos, podemos extraer las siguientes conclusiones. En primer lugar, es evidente que los informantes de nuestro corpus presentan dificultades para diferenciar el pretérito indefinido y el imperfecto, ya que los porcentajes de confusión entre ambos tiempos verbales son muy elevados. Asimismo, el tiempo de pasado más frecuente es el pretérito indefinido, y es precisamente este tiempo verbal el que consigue una mayor variedad léxica. El verbo *ser* resulta problemático en ambos tiempos verbales, presentando un elevado porcentaje de errores. Sin embargo, los verbos de movimiento, especialmente *ir*, no parecen suponer dificultades para los informantes. Por otra parte, se ha observado que el pretérito imperfecto gana variedad léxica a medida que aumenta el nivel de los estudiantes. Es decir, los del grupo G4 no solo son capaces de emplear este tiempo cometiendo menos errores que el resto de grupos sino que también utilizan más verbos distintos. En lo que respecta a los errores de conjugación del pretérito indefinido, es destacable que aparezcan en la persona *nosotros* incluso en verbos regulares.

Asimismo, con este trabajo pretendemos ilustrar una forma más de la utilidad de los corpus en la clase de español como lengua extranjera. Además de ser herramientas para seleccionar o extraer muestras de lengua real, pueden emplearse para realizar análisis sin que sea necesario poseer conocimientos de lingüística computacional. Si bien es cierto que suponen un aumento en la carga de trabajo del profesor, la utilidad de recopilar corpus de la interlengua de nuestros alumnos para su posterior análisis es manifiesta, ya que los resultados pueden destacar problemas que en principio no serían evidentes. Concluimos de esta forma nuestro trabajo, animando a todos los profesores a realizar este tipo de estudios por su interés y utilidad para la enseñanza del español como lengua extranjera.

Referencias Bibliográficas

BELLO, A. (1984): *Gramática de la lengua castellana*. Madrid: EDAF.

CARRASCO GUTIÉRREZ, A. (1999): "El tiempo verbal y la sintaxis oracional. La *consecutio temporum*", Bosque Muñoz, I. y Demonte Barreto, V. (eds.). *Gramática descriptiva de la lengua española, Vol. 2. Las construcciones sintácticas fundamentales*, 3061-3128.

GARCÍA FERNÁNDEZ, L. (2004): "El pretérito imperfecto: repaso histórico y bibliográfico", *El pretérito imperfecto,* Madrid: Gredos, 13-95.

GRANADOS, V.; MORALES, F. y DE TOMÁS, L. (1979): *Lengua española,* Madrid: Editorial Rosas.

GUTIÉRREZ ARAUS, Mª L. (1995): *Formas temporales del pasado en indicativo,* Madrid: Arco/Libros.

MORENO DE ALBA, J. G. (2006): "Valores verbales de los tiempos pasados de indicativo y su evolución", *Sintaxis histórica de la lengua española. Primer parte: La frase verbal. Vol. 1.* México, Universidad Nacional Autónoma de México, 5-94.

MUÑOZ SANZ, C. (2017): *La enseñanza del tiempo pasado del español a hablantes japoneses. Análisis contrastivo, análisis de errores y aplicación didáctica.* Tesis doctoral pendiente de publicación. Madrid: Universidad Complutense.

PORTO DAPENA, J. A. (1989): *Tiempos y formas no personales del verbo.* Madrid, Arcos/Libros.

ROJO, G. (1990): "Relaciones entre temporalidad y aspecto en el verbo español". Bosque Muñoz, I. (ed.). *Tiempo y aspecto en español,* Madrid: Cátedra, 17-44.

VEIGA, A. (2008): *"Co-pretérito" e "irreal" / "imperfecto" e "inactual". El doble valor de la forma* cantaba *en el sistema verbal español y algunos problemas conexos,* Lugo: Axac.