

## VARIABLES LINGÜÍSTICAS DE LA COMPRESIBILIDAD, EL ACENTO EXTRANJERO Y LA FLUIDEZ EN EL ESPAÑOL COMO L2. DIFERENCIAS EN FUNCIÓN DEL TIPO DE OYENTE Y DEL GRADO DE DESEMPEÑO

Enrique Santamaría Busto

New York University in Madrid  
enrique.santamaria@nyu.edu ORCID: <https://orcid.org/0000-0002-1936-9209>

**Recibido:** 10/03/21; **Aceptado:** 06/07/21; **Publicado en línea:** 28/12/21

**Citation / Cómo citar este artículo:** Enrique Santamaría Busto (2020). Variables lingüísticas de la comprensibilidad, el acento extranjero y la fluidez en el español como L2. Diferencias en función del tipo de oyente y del grado de desempeño. *Loquens*, 7(2), e075. <https://doi.org/10.3989/loquens.2020.075>

**RESUMEN:** Este estudio se propone descubrir qué variables lingüísticas se relacionan de forma más intensa con la comprensibilidad, el acento extranjero y la fluidez en el español como L2, y cómo los resultados pueden variar además en función del tipo de oyente (experto o no experto) y del grado de desempeño de los hablantes en estas dimensiones. Con este fin, 40 anglohablantes nativos que tienen el español como L2 describieron oralmente una historieta gráfica que fue posteriormente evaluada mediante escalas de 9 niveles por cuatro grupos de jueces. Los dos primeros, compuestos por 109 jueces no expertos y 42 expertos, evaluaron la comprensibilidad, el acento extranjero y la fluidez. Por su parte, los dos últimos grupos, integrados por 35 fonetistas y 35 lingüistas, se encargaron de analizar y puntuar un total de 14 variables lingüísticas (7 por grupo) relacionadas con la pronunciación, las disfluencias y los aspectos lexicogramaticales y discursivos. Los resultados revelan que la comprensibilidad se asocia a una gran variedad de elementos, mientras que el acento extranjero viene determinado especialmente por los aspectos segmentales y la fluidez por el tiempo. Asimismo, la repercusión de los parámetros lingüísticos varía en función del grado de desempeño de los hablantes, y existen diferencias entre los jueces no expertos y expertos, ya que los primeros se ven más afectados por la pronunciación y los expertos por el léxico y la gramática.

**Palabras clave:** comprensibilidad, acento extranjero, fluidez, español como L2, jueces, grados de desempeño.

**ABSTRACT:** *Linguistic variables of Comprehensibility, Accentedness and Fluency in L2 Spanish: Listener type and ability level differences.* – The aim of this study is to identify which linguistic variables are most strongly related to comprehensibility, accentedness and fluency in Spanish as L2, and how the results may also vary according to the type of listener (expert or non-expert) and to the speakers' level of ability in these three dimensions. To this end, 40 native English speakers of Spanish orally described picture narratives that were subsequently rated by four groups of raters using 9-point Likert scales. The first two groups were composed of 109 novice and 42 expert raters assessing comprehensibility, accentedness and fluency. The last two groups consisted of 35 phoneticians and 35 linguistic raters responsible for analyzing and scoring a total of 14 speech measures (7 per group) targeting pronunciation, disfluencies, lexis, grammar and discourse. The results reveal that comprehensibility is associated with a wide range of elements, while accentedness is determined especially by segmentals, and fluency by speech rate. Additionally, the impact of linguistic parameters varies according to the speakers' level of performance, and there are differences between novice and expert raters, with the former being more affected by pronunciation and the experts by lexis and grammar.

**Keywords:** comprehensibility, accentedness, fluency, L2 Spanish, raters, ability levels.

## 1. INTRODUCCIÓN

El abandono progresivo que, desde los años 90, ha sufrido la dimensión del acento extranjero en favor de otras dimensiones más importantes para la comunicación, como la inteligibilidad, la comprensibilidad o la fluidez, ha desarrollado recientemente una fértil línea de investigación sobre los aspectos lingüísticos que inciden en tales dimensiones. La inteligibilidad, o la medida en que el mensaje del hablante se comprende realmente, y la comprensibilidad, o la dificultad que experimenta el oyente para comprender dicho mensaje (*vid.* Munro y Derwing, 1995b), se insertan en lo que actualmente se conoce como «principio de inteligibilidad», siguiendo una célebre formulación planteada por Levis en el año 2005. En lo que se refiere a la pronunciación, el principio de inteligibilidad defiende que algunos elementos fónicos son más importantes que otros para el éxito de la comunicación, por lo que el objetivo de cualquier instrucción debería dirigir la atención del aprendiente hacia la producción reconocible, y no necesariamente nativa, de estas formas fonético-fonológicas.

El «principio de inteligibilidad» se contrapone a otro principio que ha marcado de forma sustancial la historia de la enseñanza-aprendizaje de la pronunciación, el «principio del acento nativo» (Levis, 2005). Este principio, que todavía goza de un gran predicamento en las aulas de segundas lenguas, tiene como objetivo la eliminación del acento extranjero, o lo que es lo mismo, la supresión de cualquier diferencia que presenta el acento de un hablante de L2 con respecto a las normas de una determinada comunidad lingüística nativa. El «principio del acento nativo» presupone así, falsamente, que todos los errores de pronunciación son igualmente importantes, que la eliminación de cualquiera de ellos redundaría directamente en una mejora de la inteligibilidad o de la comprensibilidad, y que existe un modelo de referencia (el modelo nativo) que se debe alcanzar para que el hablante pueda comunicarse con éxito.

A pesar del énfasis que tradicionalmente se ha puesto en la dimensión del acento extranjero y en los consiguientes esfuerzos por seguir y alcanzar el referente nativo, el consenso creciente entre los investigadores es que, a la hora de abordar la enseñanza de la pronunciación, los hablantes de segundas lenguas deberían ser, en primer lugar, entendidos por sus interlocutores, y a continuación que su discurso fuera fácil de comprender. Poseer un acento extranjero, en cambio, no impide ni dificulta necesariamente la comunicación (Munro y Derwing, 1995a), es una consecuencia natural del aprendizaje de una segunda lengua (Derwing y Munro, 2005), está muy vinculado a la identidad (Szabo, 2006) y es extremadamente difícil de eliminar, ya que depende de una gran diversidad de variables, entre las que destacan la edad en la que se inicia el aprendizaje de la L2 y la frecuencia de uso y de exposición (Piske *et al.*, 2001). Junto a estas diferentes perspectivas de abordar la producción oral, en los últimos años se ha suscitado también un profundo interés por otra dimensión que, a diferencia del acento extranjero, sí tiene

una profunda incidencia en la comunicación, la fluidez, la cual suele entenderse como el grado de soltura, libre de excesivas pausas y titubeos, que presenta el discurso oral del hablante (Derwing y Munro, 2015).

Dada la distinta importancia que tienen estas cuatro dimensiones para que el acto comunicativo se desarrolle con éxito, uno de los principales objetivos de los estudios ha sido proveer a los profesores de las principales claves lingüísticas en que deben focalizar su atención para promover con más eficacia la inteligibilidad, la comprensibilidad o la fluidez de sus estudiantes, y descartar las que solo son marca de acento extranjero. Para abordar este asunto, el estudio conjunto de las variables lingüísticas (tanto fónicas como no fónicas) que podrían incidir en las distintas dimensiones empezó a realizarse para el inglés como L2 (Munro y Derwing, 1995a; Trofimovich y Isaacs, 2012; Saito *et al.*, 2016), y pronto se extendió a otras lenguas como el francés (Préfontaine y Kormos, 2016; Bergeron y Trofimovich, 2017); el alemán (O'Brien, 2014); el japonés (Saito y Akiyama, 2017); el neerlandés (Pinget *et al.*, 2014); o el español (McBride, 2015; Nagle y Huensch, 2020).

Los estudios realizados hasta la fecha han demostrado que los aspectos lingüísticos que intervienen en las distintas dimensiones no son siempre los mismos e inciden con fuerzas dispares. Por ejemplo, en los últimos años se ha demostrado que la comprensibilidad es una dimensión multiparamétrica sujeta a variables léxicas, gramaticales, fónicas, discursivas y de disfluencia (Isaacs y Trofimovich, 2012), y que no todos los problemas lingüísticos relacionados con el acento extranjero tienen por qué dificultar la comprensibilidad (Trofimovich y Isaacs, 2012). En lo que concierne al acento extranjero, se ha comprobado que este depende fundamentalmente de aspectos segmentales y suprasegmentales, aunque la dificultad de la tarea puede aumentar la influencia de los problemas gramaticales y léxicos (Crowther *et al.*, 2018). Por último, en lo que respecta a la fluidez, se ha constatado que en ella intervienen aspectos temporales y marcas de disfluencia, como vacilaciones y reparaciones, aunque también pueden influir con distinta intensidad la gramática, el léxico o el ritmo (Sánchez Avendaño, 2002; Préfontaine y Kormos, 2016).

En español, la influencia de las variables lingüísticas está bien documentada en lo que concierne al acento extranjero, pero son todavía muy pocos los estudios (*vid.* Schairer, 1992; McBride, 2015; Santamaría Busto, 2015; y Nagle y Huensch, 2020) que han realizado tal análisis en una dimensión como la comprensibilidad, tan importante para la comunicación. Además, algunos de estos trabajos se han servido de tareas de expresión lectora, pero no de expresión oral. Por ejemplo, en el conocido estudio de Schairer (1992) sobre la comprensibilidad de anglohablantes de español como L2, los errores de pronunciación se cuantifican a través de una prueba de lectura, y se constata que los parámetros segmentales —especialmente las vocales—, seguidos de la resilabificación y de la articulación de la /r/, son los aspectos más influyentes. Sin embargo, el tipo de tarea impide considerar la incidencia

de otros elementos que sí podrían resultar determinantes en el habla espontánea, como el uso de las pausas, la entonación, la corrección gramatical o la precisión léxica. Por ello, el empleo de estas tareas podría provocar, en última instancia, que la importancia de algunos de los parámetros analizados estuviera sobrevalorada. Del mismo modo, en el trabajo de McBride (2015) la comprensibilidad –dentro de un estudio más global que reúne el análisis de otros aspectos– se mide a partir de la lectura de un texto, lo que impide abrir también el análisis a otras unidades lingüísticas que no pueden valorarse.

Por otra parte, los estudios que han tenido en cuenta otro tipo de pruebas más abiertas para el análisis de la comprensibilidad (Santamaría Busto, 2015; Nagle y Huensch, 2020), no han analizado todavía cómo influye el desempeño de los hablantes en esta dimensión a la hora de valorar si todos los parámetros lingüísticos tienen realmente la misma importancia, como sí se ha hecho, recientemente, en el caso del inglés (*vid.* Saito *et al.*, 2016). Dada la relevancia del español como lengua de estudio y de comunicación, la carencia de estudios sobre los parámetros que más influyen en una dimensión tan fundamental como la comprensibilidad, y si estos varían en función del grado de desempeño del hablante, resulta sorprendente. Tras el mandarín, el español es actualmente la segunda lengua en el mundo por número de hablantes nativos (Instituto Cervantes, 2020) y el tercer idioma, tras el inglés y el francés, de comunicación internacional (Constantinescu, 2015). Las proyecciones realizadas por diversos organismos, como el Instituto Cervantes (2020), indican, además, que el peso de la comunidad hispanohablante y del número de estudiantes de ELE crece a un ritmo sostenido.

Por si fuera poco, en los Estados Unidos, que es el país del que proceden los aprendientes de español analizados en esta investigación, el español es el idioma extranjero más estudiado en todos los niveles de enseñanza. Como señalan Looney y Lusin (2019), desde 1995 el español supera en volumen de estudiantes a todas las demás lenguas extranjeras juntas y en 2016 había, solo en el sistema universitario estadounidense, más de 700.000 estudiantes que aprendían español. Tal importancia tiene una explicación, y es que esta lengua es actualmente la segunda más hablada en ese país y se espera que, para 2060, los ciudadanos de origen hispano representen más de una cuarta parte de la población estadounidense, lo que convertirá a Estados Unidos en el segundo país por número de hispanohablantes en el mundo, solo precedido por México (Instituto Cervantes, 2020).

Teniendo en cuenta, por tanto, que el grueso de los estudios se ha dirigido al análisis del inglés, y que además no tengo constancia en nuestra lengua de ninguna investigación que haya considerado conjuntamente las dimensiones de la comprensibilidad, el acento extranjero y la fluidez atendiendo al grado de desempeño de los hablantes en estas dimensiones, se diseñó un experimento en el que 35 fonetistas y 35 lingüistas evaluaron respectivamente 7 variables de pronunciación y disfluencias y 7 variables lexicogramaticales y discursivas del discurso oral producido por 40 anglohablantes nativos con distintos niveles

de competencia lingüística (A2, B1, B2 y C1). Al mismo tiempo, otro grupo compuesto por 109 jueces No expertos (evaluadores sin experiencia en la enseñanza y sin gran familiaridad con el español de anglohablantes nativos) y otro formado por 42 jueces Expertos (profesores de español como L2 muy familiarizados con este tipo de interlengua) evaluaron la comprensibilidad, el acento extranjero y la fluidez de los hablantes. El empleo de dos tipos de evaluadores, No expertos y Expertos, para medir las dimensiones, se realizó con el propósito de obtener datos complementarios que reflejaran las valoraciones de los dos grupos de personas nativas más habituales que estos hablantes encuentran tanto fuera como dentro del aula. El último propósito de este trabajo es comparar entonces los resultados ofrecidos por los No expertos y Expertos, ya que el modo en el que se relacionan las dimensiones y los factores lingüísticos que intervienen en ellas pueden arrojar diferencias entre los dos grupos.

Esta investigación pretende responder, por tanto, a las siguientes preguntas:

1. ¿Qué parámetros lingüísticos inciden en la comprensibilidad, el acento extranjero y la fluidez de los hablantes cuando se contrastan las puntuaciones de los jueces No expertos y Expertos con las calificaciones numéricas que realizan especialistas de la fonética y de la lingüística?
2. ¿Qué parámetros lingüísticos permiten distinguir mejor entre los tramos con los que se puede clasificar el desempeño de los hablantes en cada dimensión?
3. ¿Se producen diferencias entre los jueces No expertos y Expertos en cuanto a los parámetros que reciben más peso a la hora de evaluar las dimensiones?

Las hipótesis de este estudio son que los parámetros lingüísticos no inciden de la misma manera en las tres dimensiones, de modo que la comprensibilidad estará abierta a una gran variedad de aspectos; que el acento extranjero dependerá fundamentalmente de los parámetros segmentales y suprasegmentales; y que la fluidez estará más influida por los aspectos temporales y las marcas de disfluencia, aunque no hay una hipótesis abierta sobre cuál de ellos será más importante. Se espera también que algunos parámetros mantengan su influencia en las tres dimensiones, aunque con distintas intensidades, y que se produzcan diferencias cuando los hablantes se segmentan por tramos de desempeño. Por último, se espera que el peso de los parámetros también difiera en función de la experiencia de los jueces.

## 2. MÉTODO

### 2.1. Participantes

#### 2.1.1. Hablantes

Los hablantes de este experimento son 40 estudiantes de Estados Unidos (17 hombres y 23 mujeres) que tienen el inglés como L1 y están cursando un semestre de sus

estudios en España, específicamente en los campus que dos universidades estadounidenses (New York University y Tufts University) tienen en Madrid. Su promedio de edad es de 20.2 años ( $D.E. = 0.61$ ,  $rango = 19-22$ ) y han sido clasificados por sus universidades en distintos niveles de competencia lingüística (A2, B1, B2 y C1;  $n = 10$  por grupo). Los hablantes llevan estudiando español un promedio de 7.74 años ( $D.E. = 3.16$ ,  $rango = 2-15$ ), y remiten un porcentaje medio de uso del español de 52.38 % ( $D.E. = 23.42$ ,  $rango = 10-90$ ). Su grado de contacto en español con hispanohablantes<sup>1</sup>, sobre una escala Likert de 9 niveles donde 1 representa «Ningún contacto» y 9 «Mucho contacto» es de 5.15 puntos ( $D.E. = 1.92$ ,  $rango = 2-9$ ).

### 2.1.2. Evaluadores

Los cuatro grupos de evaluadores se obtuvieron a través de llamamientos personales por correo electrónico, anuncios en plataformas de Internet y, en el caso también de los profesores y de los lingüistas, de llamamientos realizados tanto por parte de este investigador como por parte de directores o coordinadores de estudios que trabajan para universidades estadounidenses en España. Por su parte, el grupo de fonetistas estaba integrado fundamentalmente por antiguos alumnos del –lamentablemente– ahora extinto Máster de Fonética y Fonología del CSIC y la UIMP que se cursaba en la ciudad de Madrid.

#### 2.1.2.1. Evaluadores No expertos y Expertos

Estos dos grupos de evaluadores están compuestos por 109 jueces No expertos y 42 Expertos. El grupo de No expertos tiene un promedio de edad de 41.44 años ( $D.E. = 12.60$ ,  $rango = 20-66$ ), y el de Expertos de 45.95 años ( $D.E. = 8.70$ ,  $rango = 26-65$ ). Ninguno de los jueces No expertos tiene experiencia en la enseñanza del español como L2/LE, ni estudios en filología o lingüística a nivel universitario. Los jueces Expertos trabajan, en cambio, como profesores de español como L2 y tienen 19.52 años de experiencia ( $D.E. = 9.16$ ,  $rango = 3-44$ ). Todos los jueces No expertos y Expertos residían en España, fundamentalmente en la ciudad de Madrid. Debido a que el conocimiento de la L1 del hablante es un factor que puede influir en las evaluaciones (Winke *et al.*, 2013; Bergeron y Trofimovich, 2017), y a que en España la enseñanza de esta lengua es la que predomina con claridad en el sistema educativo obligatorio (Muñoz Zayas, 2013) y, por tanto, un porcentaje creciente y cada vez más mayoritario de la población la habla (INE, 2017), en este estudio solo participaron jueces que tuvieron conocimientos de la lengua inglesa.

<sup>1</sup> Se incluyeron las preguntas de «Uso del español» y «Contacto en español con hispanohablantes nativos» porque algunos estudiantes no solo emplean el español con hispanófonos, sino que a veces también lo utilizan entre ellos.

Asimismo, se buscó que el grado de familiaridad de los jueces No expertos con el español de los hablantes no fuera el mismo que el que se esperaba de los jueces Expertos. Con este fin, para poder realizar el experimento los participantes debían contestar a unas preguntas de control al comienzo del cuestionario en las que se estableció un límite de 4 puntos sobre 9 tanto en grado de familiaridad con el español de anglohablantes nativos como en grado de contacto en español con estos hablantes. Como puede observarse en la Tabla 1, los jueces No expertos presentan un escaso grado de familiaridad con el español de este tipo de hablantes ( $M = 2.37$ ,  $D.E. = 1.32$ ) y poco contacto en español ( $M = 2.05$ ,  $D.E. = 1.28$ ). Por el contrario, los jueces Expertos manifiestan mucha familiaridad ( $M = 8.74$ ,  $D.E. = 0.45$ ) y mucho contacto en español ( $M = 8.52$ ,  $D.E. = 0.80$ ). Aunque con cifras sensiblemente inferiores a las de estos resultados, los Expertos señalan también un mayor contacto con otros hablantes que tienen el español como L2.

#### 2.1.2.2. Evaluadores fonetistas y lingüistas

Estos dos grupos de evaluadores están formados por 35 lingüistas (profesores de E/L2 distintos a los anteriores) que evalúan 7 parámetros lexicogramaticales y discursivos y 35 fonetistas que evalúan 7 parámetros de pronunciación y de disfluencia. Como se observa en la Tabla 1, el promedio de edad de los jueces fonetistas es de 32.14 años ( $D.E. = 5.67$ ,  $rango = 24-50$ ), y el de los profesores de 45.09 ( $D.E. = 11.20$ ,  $rango = 26-67$ ). Los dos grupos están altamente especializados en sus respectivas disciplinas. El 97.1 % de los fonetistas tiene un Máster en fonética y fonología, en lingüística o un doctorado en este campo, el 51.4 % ha publicado sobre esta materia y el 54.3 % ha dado clases de estos contenidos. Por su parte, el 74.3 % de los profesores de español tiene un Máster o un doctorado en filología, lingüística o enseñanza del español como L2/LE, y el 25.7 % restante una licenciatura en filología hispánica, traducción e interpretación, o filología inglesa. El 65.7 % tiene también formación en evaluación (los fonetistas, un 11.4 %). El 100 % de los fonetistas y de los lingüistas habla inglés.

Por otra parte, todos los lingüistas tienen experiencia en la enseñanza del español como L2/LE, con un promedio de 16.83 años ( $D.E. = 9.64$ ,  $rango = 3-40$ ). El 34.3 % de los fonetistas señala que también tiene este tipo de experiencia, aunque en su caso esta es mucho menor, casi siempre inferior a un año. Respecto al grado de familiaridad con el español de anglohablantes y el grado de contacto que mantienen con ellos en español, los profesores señalan respectivamente en una escala Likert de 9 niveles un 8.34 ( $D.E. = 1.00$ ) de promedio en familiaridad (donde 1 = Ninguna familiaridad, y 9 = Muy familiarizado), y un 8.23 ( $D.E. = 1.11$ ) en contacto (1 = Ningún contacto, 9 = Mucho contacto). Los fonetistas señalan un grado de familiaridad de 4.83 ( $D.E. = 1.87$ ), y un grado de contacto en español de 3.86 ( $D.E. = 2.26$ ).

Tabla 1: Características de los jueces (N = 221).

VARIABLES	No expertos (n=109)	Expertos (n=42)	Fonetistas (n=35)	Lingüistas (n=35)
<b>SEXO</b>				
<i>Hombre</i>	44.0% (48)	26.2% (11)	28.6% (10)	31.4% (11)
<i>Mujer</i>	56.0% (61)	73.8% (31)	71.4% (25)	68.6% (24)
<b>EDAD</b>				
<i>Media (D.E.)</i>	41.44 (12.60)	45.95 (8.70)	32.14 (5.67)	45.09 (11.20)
<i>Mediana</i>	41	45	31	45
<b>FAMILIARIDAD con el español de anglohablantes<sup>a</sup></b>				
<i>Media (D.E.)</i>	2.37 (1.32)	8.74 (0.45)	4.83 (1.87)	8.34 (1.00)
<i>Mediana</i>	2.00	9.00	5.00	9.00
<b>CONTACTO en español con anglohablantes<sup>b</sup></b>				
<i>Media (D.E.)</i>	2.05 (1.00)	8.52 (0.80)	3.86 (2.26)	8.23 (1.11)
<i>Mediana</i>	1.00	9.00	4.00	9.00
<b>CONTACTO en español con otros hablantes de español L2<sup>b</sup></b>				
<i>Media (D.E.)</i>	2.20 (2.29)	5.17 (3.34)	4.80 (3.21)	5.94 (3.27)
<i>Mediana</i>	1.00	5.00	4.80 (3.21)	5.94 (3.27)
<b>AÑOS de experiencia como profesor de español L2/LE</b>				
<i>Media (D.E.)</i>	--	19.52 (9.16)	--	16.83 (9.64)
<i>Mediana</i>		20		18

<sup>a</sup> 1=Ninguna familiaridad, 9 = Muy familiarizado; <sup>b</sup> 1=Ningún contacto en español, 9=Mucho contacto en español.

## 2.2. Procedimiento

### 2.2.1. Tarea empleada

Siguiendo el tipo de tarea utilizado por otros investigadores, la tarea utilizada para obtener las muestras orales consistió en la descripción de una historia dispuesta en viñetas. Dicha historieta, propuesta por Celce-Murcia *et al.* (2010: 315), está estructurada en cuatro viñetas ordenadas que muestran a un monitor de acampada haciendo senderismo por el campo con un grupo de chicos jóvenes. Durante el recorrido, el hombre se da cuenta de que los chicos tienen miedo de estar en ese lugar, de modo que, para gastarles una broma, cuando llega la noche les cuenta una historia de miedo que les infunda aún más temor. La última viñeta finaliza con los chicos durmiendo plácidamente, mientras que el monitor no concilia el sueño porque se ha creído su propia historia.

La elección de esta tarea se debe a varias razones. En primer lugar, sus características hacen que sea más compleja, ya que, al contrario que otras, como el monólogo o la lectura de un texto, esta tarea obliga a emplear el vocabulario que requiere exactamente la historia para describirla con éxito, lo que impide que los hablantes puedan reconducir el discurso hacia usos lingüísticos en los que se sienten más cómodos. En segundo lugar, la tarea

reproduce también situaciones cotidianas que los estudiantes pueden encontrar fuera del aula, como es la descripción de lo que sucede en una historia atendiendo a los objetos, los personajes y sus acciones. En tercer lugar, se espera que una tarea cognitivamente más exigente discrimine mejor entre tramos de desempeño de los hablantes que una tarea que no lo es, ya que estimulan la producción de versiones lingüísticas más complejas. Y, por último, se tiene en cuenta que este tipo de tarea ha sido –y sigue siendo– el más utilizado en las investigaciones sobre evaluación de la pronunciación de una L2, lo que facilita la comparación de los resultados que se obtengan.

Todas las muestras orales se recogieron en una habitación silenciosa mediante una grabadora H1 Handy Recorder en formato .wav, a una frecuencia de muestreo de 44.1 kHz y una resolución de 16 bits. Los hablantes dispusieron de aproximadamente 30-45 segundos de preparación para que pudieran familiarizarse con la historieta y pudieran plantear preguntas de carácter técnico (pero no lingüístico) si no entendían algo. Este tiempo de preparación es el mismo que el proporcionado en otros estudios para tareas similares (Derwing *et al.*, 2004, 2006; Bergeron y Trofimovich, 2017), y se consideró apropiado porque permitía a los hablantes entender la historieta sin monitorizar el discurso hasta un punto en que afectara la espontaneidad de su producción oral (Ortega, 1999).

En línea con otras investigaciones previas que demuestran cómo los jueces pueden realizar evaluaciones fiables de estas dimensiones partiendo de muestras relativamente cortas, de entre 20 y 30 segundos (Trofimovich y Isaacs, 2012), las 40 muestras evaluadas en el presente experimento abarcan extractos de aproximadamente 30 segundos (*rango* = 27-37), correspondientes al comienzo de la historia. Como es habitual en este tipo de estudios, los falsos comienzos y los posibles carraspeos situados al inicio se eliminaron, y los finales de todas las muestras coincidieron con pausas naturales.

### 2.2.2. Evaluación de las muestras

Debido a la duración de los cuestionarios y a las dificultades encontradas para obtener un número significativo de jueces que pudiera realizar el experimento de forma presencial en una misma sesión, la evaluación de las muestras se realizó en Internet a través de la plataforma *Qualtrics*. El experimento se podía hacer en varias sesiones, aunque una vez iniciado este solo permanecía activo durante un plazo de 72 horas. Asimismo, antes de proceder a la evaluación de las 40 muestras los jueces debían puntuar tres muestras adicionales que servían de familiarización y práctica, y podían consultar permanentemente la historietta pulsando sobre un enlace situado en la parte superior de la página. Por otra parte, se informó a los evaluadores que los hablantes tenían diferentes niveles de competencia lingüística y que los aspectos evaluados debían puntuarse de forma independiente, sin establecer correspondencias entre ellos o con los posibles niveles de lengua. También se les pidió que intentaran utilizar todas las posiciones de las escalas. Las muestras se presentaron de forma aleatoria para cada participante, de modo que todos los jueces las evaluaron en un orden distinto.

#### 2.2.2.1. Evaluación de la comprensibilidad, el acento extranjero y la fluidez

En el caso de los jueces No expertos y Expertos la evaluación de las tres dimensiones se realizó de forma conjunta para cada muestra, ya que no hay diferencias entre evaluarlas juntas de forma simultánea o de forma consecutiva por separado (O'Brien, 2016). Tomando como referencia otros estudios que consideran que la medición de estas dimensiones se ajusta a las primeras impresiones e intuiciones del oyente (Trofimovich y Isaacs, 2012), las muestras solo podían oírse una vez, y además solo podían evaluarse una vez terminado el audio, de modo que el oyente debía escuchar toda la muestra antes de poder puntuarla. Las tres dimensiones evaluadas por estos jueces se definieron y se midieron a través de escalas Likert de la siguiente manera:

*Comprensibilidad*: grado de dificultad que experimenta el oyente para entender lo que dice el hablante. Se mide con una escala donde 1 = Muy difícil de entender; y 9 = Muy fácil de entender.

*Acento extranjero*: grado de diferencia que presenta el acento de un hablante con respecto a un nativo de español.

Se mide con una escala donde 1 = Fuerte acento extranjero; y 9 = Sin acento extranjero.

*Fluidez*: grado de soltura, libre de excesivas pausas y titubeos, que presenta el discurso oral del hablante. Se mide con una escala donde 1 = Muy poca fluidez; y 9 = Excelente fluidez.

El experimento tenía una duración aproximada de 45 minutos, y los evaluadores podían realizar comentarios.

#### 2.2.2.2. Evaluación de los 14 parámetros lingüísticos

Los jueces fonetistas y lingüistas debían evaluar con una escala Likert de 9 niveles 7 aspectos lingüísticos distintos que específicamente se correspondían con 7 aspectos de «Pronunciación y disfluencia» (grupo de fonetistas) y 7 aspectos de «Lexicogramática y discurso» (grupo de lingüistas). Los 7 parámetros evaluados por los fonetistas son:

1. *Vocales y consonantes*: grado de interferencia fonético-fonológica de la L1 del hablante a la hora de pronunciar los segmentos del español.
2. *Acento*: grado de interferencia fonético-fonológica de la L1 del hablante a la hora de realizar y colocar el acento en español.
3. *Ritmo*: grado de interferencia fonético-fonológica de la L1 del hablante a la hora de marcar la duración silábica, el timbre de la vocal, la configuración de grupos tonales y la tensión articulatoria.
4. *Entonación*: grado de interferencia fonético-fonológica de la L1 del hablante a la hora de realizar la melodía tonal de un enunciado.

Estos cuatro parámetros se miden con una escala donde 1 = Mucha interferencia del inglés; y 9 = Sin interferencia del inglés.

5. *Tempo*: grado de velocidad de elocución del hablante a la hora de realizar el enunciado. Se mide con una escala donde 1 = Demasiado lenta/rápida; y 9 = Óptima.
6. *Pausas*: uso que hace el hablante de las pausas durante la producción de su mensaje. Se mide con una escala donde 1 = Muy poco natural; y 9 = Muy natural.
7. *Repeticiones, reparaciones y falsos comienzos*: frecuencia con la que el hablante realiza repeticiones, reparaciones y falsos comienzos durante la producción de su mensaje. Se mide con una escala donde 1 = Muy frecuentes; y 9 = Muy infrecuentes.

Debido a que la evaluación de estos aspectos requiere un análisis profundo de la muestra, los jueces tenían la opción de volver a escuchar el audio las veces que quisieran. Los 7 parámetros lingüísticos se evaluaron juntos en cada muestra, dispuestos en el orden en que se han presentado anteriormente, y se podían realizar comentarios. El experimento tenía una duración aproximada de 120 minutos, aunque varios jueces señalaron que les había llevado más tiempo realizarlo.

Por su parte, los jueces lingüistas evaluaron 7 aspectos relacionados con el léxico, la gramática y el discurso. En línea con otros estudios (Bergeron y Trofimovich, 2017; Saito, 2019), para minimizar la influencia de la pronunciación en la evaluación de estos aspectos, este grupo de jueces evaluó solo la transcripción ortográfica de las muestras con algunas pequeñas modificaciones: así, se eliminaron la escritura de pausas (por ejemplo *ah*, *am*), los errores específicos de pronunciación (p. ej. «cuatrou» o «líder» se escribieron como «cuatro» y «líder»)<sup>2</sup>, y la puntuación, en este último caso con el fin de evitar la interpretación del propio investigador a la hora de situar las pausas o de valorar la entonación.

Al igual que el grupo de fonetistas, los jueces lingüistas tuvieron la opción de evaluar las muestras sin límite de tiempo. En su caso, se calculó una duración aproximada del experimento algo menor, de aproximadamente 100 minutos. Los parámetros lingüísticos que tuvieron que evaluar se dispusieron siempre en el mismo orden y en cada muestra se ofrecía también la posibilidad de realizar comentarios. Los parámetros evaluados son los siguientes:

1. *Riqueza léxica*: grado de amplitud y variedad de vocabulario que presenta el hablante en español a la hora de describir el comienzo de la historieta. Se mide con una escala donde 1 = Vocabulario simple y repetitivo; 9 = Vocabulario variado.
2. *Precisión léxica*: grado de pertinencia léxica que presenta el hablante en español a la hora de describir el comienzo de la historieta. Se mide con una escala donde 1 = Selección inapropiada de palabras; 9 = Selección apropiada de palabras.
3. *Complejidad sintáctica*: grado de producción de estructuras sintácticas complejas que presenta el hablante en español a la hora de describir el comienzo de la historieta. Se mide con una escala donde 1 = Sintaxis simple y fragmentada; 9 = Sintaxis elaborada.
4. *Corrección gramatical*: grado de presencia de errores de tipo morfológico y sintáctico que introduce el hablante en español a la hora de describir el comienzo de la historieta. Se mide con una escala donde 1 = Muchos errores gramaticales; 9 = Sin errores gramaticales.
5. *Coherencia discursiva*: grado de sentido lógico que presentan y mantienen entre sí los elementos del texto. Se mide con una escala donde 1 = Narración sin sentido lógico, ideas no relacionadas; 9 = Narración con sentido lógico, ideas relacionadas.

<sup>2</sup> En los casos de centralización vocálica en los que no quedaba claro si había un problema de interferencia fonética o de error gramatical («historia peligrosa» o «peligrosa», por ejemplo) se optó por no corregir el error, ya que estos fallos de concordancia aparecen con frecuencia no solo en los discursos orales de estos hablantes (Grüter *et al.*, 2012), sino también escritos (Madrid, 1999; Fleissig, 2009), donde hay más margen de monitorización del discurso.

6. *Cohesión discursiva*: grado de presencia y uso de procedimientos de expresión que ayudan a conectar y relacionar elementos internos de la acción. Se mide con una escala donde 1 = No hay conectores y elementos de cohesión; 9 = Conectores y elementos de cohesión en lugares adecuados.
7. *Profundidad discursiva*: grado de elaboración de la información que presenta el hablante a la hora de describir el comienzo de la historieta. Se mide por tanto con una escala donde 1 = Narración simple, sin detalles; 9 = Narración elaborada, detallada.

### 3. RESULTADOS

#### 3.1. Incidencia de los parámetros lingüísticos en las dimensiones

Este apartado tiene como objetivo analizar la influencia que ejercen los parámetros léxicos, gramaticales, discursivos, fónicos y de disfluencia en la comprensibilidad, el acento extranjero y la fluidez de los hablantes. Los análisis de la fiabilidad alcanzada por todos los jueces, mediante alfa de Cronbach, registra coeficientes muy elevados, tanto en la evaluación de las tres dimensiones como de los 14 parámetros lingüísticos ( $\alpha \geq 0.95$ ).

Con el fin de comprobar si las puntuaciones de No expertos y Expertos pueden representarse juntas en un único valor, se analiza a continuación la existencia de posibles diferencias en la evaluación de la comprensibilidad, el acento extranjero y la fluidez. Para ello se realizaron tres *t* de Student independientes ( $p > 0.05$  en los tests K-S para No expertos y S-W para Expertos) con cálculos del tamaño del efecto mediante *g* de Hedges o *d* de Cohen corregida. Los resultados revelan que los jueces Expertos ofrecen siempre puntuaciones más altas, y que estas son estadísticamente significativas en comprensibilidad ( $t(149) = 3.13$ ,  $p = 0.002$ ,  $g = 0.58$ ) y en acento extranjero ( $t(149) = 3.83$ ,  $p < 0.001$ ,  $g = 0.70$ ), aunque no en fluidez ( $t(149) = 1.60$ ,  $p = 0.112$ ,  $g = 0.29$ ). Vista la existencia de diferencias significativas con tamaños del efecto moderados en dos de estas tres dimensiones, los análisis en los que intervienen estos jueces se mantienen siempre separados.

A continuación, se procedió a someter las 14 variables evaluadas por los jueces fonetistas y lingüistas a un análisis factorial con el objetivo de determinar si es posible sustituir los 14 parámetros lingüísticos por un número reducido de factores que explique la mayoría de la varianza observada. Para el cálculo del análisis factorial se empleó el Análisis de Componentes Principales y se implementó el método ortogonal de rotación Varimax con normalización Kaiser con el fin de maximizar las ponderaciones de los factores extraídos (Tabachnick y Fidell, 2007: 620-621). Como puede observarse en la Tabla 2, el análisis revela la existencia de dos factores subyacentes a las 14 variables empíricas (Bartlett =  $\chi^2_{291} = 1257.41$ ,  $p < 0.001$ ; KMO = 0.896), que suman un alto 93.5 % de variabilidad total explicada. Por otro lado, ambos factores están bastante equilibrados en importancia, ya que el

Tabla 2: Análisis factorial exploratorio. Factores lingüísticos.

MÉTODO: Componentes Principales con rotación Varimax				Factor 1	Factor 2
CONDICIONES: KMO = .896		% Varianza total explicada		52.66 %	40.84 %
Bartlett: $p < .001$		% Acumulado de Varianza		52.66 %	93.50 %
Descriptivos				Saturaciones >.500	
Variable empírica	Media	D.E.	Comunalidad	Factor 1 (Lex. Gram. y Disc.)	Factor 2 (Pron. y Disf.)
<i>Precisión léxica</i>	4.53	1.51	.944	.914	-
<i>Complejidad sintáctica</i>	4.61	1.54	.985	.910	-
<i>Cohesión discursiva</i>	4.53	1.48	.985	.909	-
<i>Profundidad</i>	4.75	1.60	.980	.907	-
<i>Coherencia discursiva</i>	5.25	1.46	.971	.907	-
<i>Riqueza léxica</i>	4.94	1.50	.975	.899	-
<i>Corrección gramatical</i>	4.43	1.68	.953	.885	-
<i>Vocales y consonantes</i>	4.47	1.33	.887	-	.917
<i>Entonación</i>	4.56	1.09	.956	-	.905
<i>Acento</i>	4.71	1.07	.942	-	.883
<i>Ritmo</i>	4.56	1.15	.961	-	.860
<i>Tempo</i>	5.13	1.44	.852	-	.742
<i>Uso de la pausas</i>	4.66	1.33	.871	-	.736
<i>Repet. Repar. y Fal. Com.</i>	5.37	1.41	.827	-	.618

factor 1 y el factor 2 explican respectivamente un 52.7 % y un 40.8 % de este total de la varianza.

Si se observan las variables que definen a cada uno de los factores, se comprueba que el factor 1 está formado por las siete variables evaluadas por el grupo de lingüistas, de modo que se recogerán con el nombre de «Lexicogramática y discurso». Por su parte, el segundo factor suma la información de las siete variables empíricas evaluadas por los jueces fonetistas, de modo que se denominará «Pronunciación y disfluencias». Con el fin de recoger estas dos nuevas variables estadísticas en el modelo de regresión se calculó a continuación una media ponderada a partir de los valores de las cargas factoriales de los 14 parámetros en su correspondiente factor. Este promedio se obtiene multiplicando las cargas factoriales de los parámetros por sus puntuaciones reales obtenidas en la escala Likert y dividiendo a continuación la suma de estos elementos entre la suma de sus pesos ponderales. Siguiendo las recomendaciones de Pardo y San Martín (2010: 399), se comprobó el cumplimiento de los supuestos de homocedasticidad, linealidad, normalidad, no colinealidad e independencia sobre todas las variables.

En este análisis el estudio de la no colinealidad arrojó coeficientes de correlación superiores a 0.700 entre los factores ( $r = 0.788$ ) y, en la mayoría de los

casos, entre los factores y las dimensiones. Para algunos investigadores, estos coeficientes de correlación introducen un problema de colinealidad que desaconseja que se consideren juntas las variables en el mismo

análisis, dada la dificultad que tiene el modelo para evaluar las aportaciones de cada predictor sin que se superpongan entre ellos (Allison, 1999: 149; Field, 2009: 224; Pardo y San Martín, 2010: 401-402). Ante esta situación, algunas de las soluciones típicas sugeridas por estos y otros autores consisten, bien en eliminar una de las variables independientes que causan el problema o bien hacer una media entre ellas (Allison, 1999: 147; Tabachnick y Fidell, 2007: 91). Teniendo en cuenta que solo se dispone de dos variables predictoras y que no se quiere prescindir de ninguna de ellas, se realizaron dos modelos independientes de regresión univariada por cada uno de estos factores ( $X_1$ - $Y$ ;  $X_2$ - $Y$ ). Esta estrategia consigue así eliminar del modelo la presencia de una variable parcialmente repetida sobre otra, sin privilegiar el estudio de solo una de ellas. Para facilitar la comparación de los resultados y simplificar su presentación, los análisis se exponen juntos en una misma tabla y en ella se incluyen tanto los obtenidos en el grupo de No expertos como en el de Expertos sobre el total de la muestra.

La Tabla 3 presenta los resultados obtenidos sobre la muestra total de hablantes en las tres dimensiones. Como puede observarse, la proporción de variabilidad que puede explicar cada variable predictora en la comprensibilidad es muy grande (cerca del 80 %,  $p < 0.001$ ), aunque se pueden apreciar diferencias entre jueces No expertos y Expertos. Para los primeros, el factor de «Pronunciación y disfluencias» tiene una mayor capacidad predictiva de la comprensibilidad, mientras que para los Expertos es el

**Tabla 3:** Análisis de Regresión Lineal Simple. Coeficientes de regresión, tests *t*, y tamaños del efecto usando los factores de «Pronunciación y disfluencias» y «Lexicogramática y discurso» como predictores lingüísticos de la comprensibilidad, el acento extranjero y la fluidez de los hablantes. Comparativa en función de la experiencia de los jueces.

Predictor lingüístico	No expertos				Expertos			
	R <sup>2</sup> ajustado	Coef. B de regresión	Test <i>t</i>		R <sup>2</sup> ajustado	Coef. B de regresión	Test <i>t</i>	
			/Valor/	valor <i>p</i>			/Valor/	valor <i>p</i>
COMPRESIBILIDAD								
<b>LEXICOGR. y DISCURSO</b>	.812	0.848	13.03**	.000	.794	0.766	12.30**	.000
<b>PRONUNC. y DISE.</b>	.832	1.120	13.94**	.000	.786	0.994	12.00**	.000
ACENTO EXTRANJERO								
<b>LEXICOGR. y DISCURSO</b>	.543	0.595	6.71**	.000	.417	0.546	5.37**	.000
<b>PRONUNC. y DISE.</b>	.926	1.015	22.08**	.000	.874	1.016	16.49**	.000
FLUIDEZ								
<b>LEXICOGR. y DISCURSO</b>	.782	0.932	11.88**	.000	.763	0.914	11.24**	.000
<b>PRONUNC. y DISE.</b>	.842	1.257	14.23**	.000	.826	1.239	13.62**	.000

\*\**p* < 0.01.

factor de «Lexicogramática y discurso». En lo que concierne al acento extranjero, el factor de «Pronunciación y disfluencias» tiene mucha mayor capacidad predictiva sobre esta dimensión (aproximadamente el 90 %), frente al factor de «Lexicogramática y discurso», que tiene un peso inferior al 55 %. En los dos casos, el grupo de No expertos ofrece valores ligeramente más elevados que el de Expertos, y las regresiones resultan significativas con *p* < 0.001. En el caso de la fluidez, el peso de la variable «Pronunciación y disfluencias» es ligeramente más alto que el de «Lexicogramática y discurso» (en los dos grupos por encima del 82 % para el primero e inferior al 77 % para el segundo, todos con *p* < 0.001). Si se comparan los dos grupos de jueces los efectos son muy similares en ambos, con una leve diferencia también a favor de los No expertos.

A continuación se analiza, mediante correlaciones parciales, la asociación que presentan las 7 variables de «Pronunciación y disfluencias» por un lado, y las 7 de «Lexicogramática y discurso» por el otro, con las dimensiones de comprensibilidad, acento extranjero y fluidez considerando los dos grupos de evaluadores sobre el total de la muestra. Con este fin, se calcula en primer lugar el grado de asociación entre las 7 variables de «Pronunciación y disfluencias» con las tres dimensiones, mientras se controla la influencia de las variables de «Lexicogramática y discurso». Como muestra la Tabla 4, los valores de las 7 variables de «Pronunciación y disfluencias» son por lo general altos, aunque la intensidad de los coeficientes y el orden con que se clasifican varía entre las tres dimensiones y los dos grupos de jueces.

Tanto para los No expertos como para los Expertos las variables de «Tempo» y «Pausas» son las que muestran una asociación más fuerte con la comprensibilidad (*r* ≥ 0.62) y con la fluidez (*r* ≥ 0.83), mientras que los aspectos segmentales son los que menos lo están (*r* ≤ 0.53). Por el contrario, en los dos grupos las «Vocales y consonantes», el «Ritmo», la «Entonación»

y el «Acento» son los que mantienen la mayor asociación con el acento extranjero (*r* ≥ 0.86). Esto es así especialmente en «Vocales y consonantes» (*r* = 0.95 para los No expertos y *r* = 0.93 para los Expertos), seguidas del «Ritmo» (*r* = 0.92 para los No expertos y *r* = 0.87 para los Expertos), y la «Entonación» (*r* = 0.92 para los No expertos y *r* = 0.86 para los Expertos).

Con el mismo objetivo de observar ahora cómo se relacionan las 7 variables de «Lexicogramática y discurso», se realizaron nuevas correlaciones parciales mientras se controlaba la influencia de las variables de «Pronunciación y disfluencias». Como puede observarse, la intensidad de los coeficientes y el orden con que se clasifican las variables varía también entre las tres dimensiones y los dos grupos de jueces. De modo similar a lo que ocurría con las variables de «Pronunciación y disfluencias», de nuevo aquí los valores de los No expertos suelen ser iguales o ligeramente superiores a los de los Expertos.

Para los No expertos, la «Precisión léxica» (*r* = 0.67), seguida de «Riqueza léxica» (*r* = 0.65), son las variables que ofrecen la mayor correlación con la comprensibilidad, mientras que para los Expertos estas son la «Corrección gramatical» (*r* = 0.68) y la «Precisión léxica» (*r* = 0.65). Para el acento extranjero, ninguna de estas variables resulta significativa en los dos grupos de jueces. En lo que concierne a la fluidez, las variables que obtienen correlaciones más altas en el caso de los jueces No expertos son la «Riqueza léxica» y la «Precisión léxica» (ambas con *r* = 0.72), y para los Expertos la «Riqueza léxica» (*r* = 0.67) y la «Profundidad discursiva» (*r* = 0.66).

### 3.2. La comprensibilidad, el acento extranjero y la fluidez en función de sus tramos de desempeño

Este apartado tiene como objetivo analizar cómo los 14 parámetros lingüísticos permiten diferenciar entre tramos de desempeño en comprensibilidad, acento

**Tabla 4:** Correlaciones parciales entre las 7 variables de «Pronunciación y disfluencias» y las 7 variables de «Lexicogramática y discurso» con la comprensibilidad, el acento extranjero y la fluidez en el grupo de No expertos y de Expertos. Subrayadas, las tres variables más importantes para cada dimensión.

	No Expertos			Expertos		
	Comprensibilidad	Acento	Fluidez	Comprensibilidad	Acento	Fluidez
<i>Voc. y Conson.<sup>a</sup></i>	0.58*	<u>0.95*</u>	0.53*	0.52*	<u>0.93*</u>	0.53*
<i>Acento<sup>a</sup></i>	0.58*	0.90*	0.56*	0.47*	<u>0.86*</u>	0.54*
<i>Ritmo<sup>a</sup></i>	<u>0.73*</u>	<u>0.92*</u>	0.70*	0.61*	<u>0.87*</u>	0.69*
<i>Entonación<sup>a</sup></i>	0.64*	<u>0.92*</u>	0.62*	0.53*	<u>0.86*</u>	0.60*
<i>Tempo<sup>a</sup></i>	<u>0.79*</u>	0.70*	<u>0.88*</u>	<u>0.73*</u>	0.69*	<u>0.87*</u>
<i>Pausas<sup>a</sup></i>	<u>0.73*</u>	0.70*	<u>0.87*</u>	0.62*	0.65*	<u>0.83*</u>
<i>Rep. Repa. y FF.CC.<sup>a</sup></i>	0.64*	0.46	<u>0.75*</u>	0.63*	0.44	<u>0.70*</u>
<i>Riqueza Léxica<sup>b</sup></i>	0.65*	0.28	0.72*	0.62*	0.17	0.67*
<i>Precisión Léxica<sup>b</sup></i>	0.67*	0.42	0.72*	<u>0.65*</u>	0.12	0.65*
<i>Complej. Sint.<sup>b</sup></i>	0.58*	0.32	0.64*	0.59*	0.15	0.62*
<i>Correc. Gram.<sup>b</sup></i>	0.64*	0.46	0.59*	<u>0.68*</u>	0.14	0.56*
<i>Coherencia<sup>b</sup></i>	0.64*	0.31	0.64*	0.63*	0.16	0.60*
<i>Cohesión<sup>b</sup></i>	0.58*	0.30	0.65*	0.58*	0.13	0.63*
<i>Profundidad<sup>b</sup></i>	0.55*	0.24	0.68*	0.53*	0.20	0.66*

<sup>a</sup>Las variables de «Lexicogramática y discurso» (señaladas con el superíndice<sup>b</sup>) están apartadas de cada correlación.

<sup>b</sup>Las variables de «Pronunciación y disfluencias» (señaladas con el superíndice<sup>a</sup>) están apartadas de cada correlación.

\* $\alpha < 0.007$  (corrección de Bonferroni).

**Tabla 5:** ANOVAS de 1 factor: resumen de las diferencias entre medias (desviaciones estándar) de los cuatro tramos de competencia en cada una de las tres dimensiones. Valoraciones obtenidas a partir del promedio por hablante según No expertos y Expertos.

Niveles	No expertos			Expertos		
	Compren. <sup>1</sup>	Acento <sup>2</sup>	Fluidez <sup>1</sup>	Compren. <sup>2</sup>	Acento <sup>1</sup>	Fluidez <sup>1</sup>
<i>Inic.<sup>a</sup></i>	3.83 (0.41)	2.80 (0.41)	3.15 (0.51)	4.55 (0.62)	3.34 (0.67)	3.43 (0.56)
<i>Int.<sup>b</sup></i>	5.28 (0.53)	3.90 (0.27)	4.81 (0.52)	5.88 (0.32)	4.68 (0.31)	5.06 (0.54)
<i>Av.<sup>c</sup></i>	6.37 (0.33)	4.75 (0.26)	6.06 (0.31)	6.79 (0.39)	5.47 (0.17)	6.31 (0.44)
<i>Sup.</i>	7.48 (0.38)	5.96 (0.49)	7.24 (0.40)	7.86 (0.38)	6.56 (0.47)	7.45 (0.34)

<sup>a</sup>Las celdas con trama indican diferencias significativas al 5 % entre el tramo Inicial y otros niveles.

<sup>b</sup>El relleno de color amarillo indica diferencias significativas al 5 % entre el tramo Intermedio y los tramos Avanzado o Superior.

<sup>c</sup>El relleno de color azul indica diferencias significativas al 5 % entre los tramos Avanzado y Superior.

<sup>1</sup>Comparaciones con corrección de Bonferroni para homogeneidad de varianzas ( $p > 0.05$  en la prueba de Levene).

<sup>2</sup>Comparaciones con corrección de T3 de Dunnett para heterocedasticidad de varianzas ( $p < 0.05$  en la prueba de Levene).

extranjero y fluidez. Con este fin, los 40 hablantes se clasificaron primero en cuatro tramos de desempeño (Inicial, Intermedio, Avanzado y Superior;  $n = 10$  por grupo) según las puntuaciones que otorgaron los jueces No expertos y Expertos (véase la Tabla 5). El tramo Inicial recoge aquí, por tanto, a los diez hablantes con peor desempeño evaluados por cada grupo de jueces, el tramo Intermedio a los diez siguientes, y así sucesivamente hasta llegar al tramo Superior donde se sitúan los diez mejores. Para estudiar el posible efecto de interacción entre los dos grupos de evaluadores y los cuatro tramos de desempeño en cada una de las dimensiones se realizan tres ANOVAS de dos factores ( $p > 0.05$  en el test S-W). Los resultados de estos tres ANOVAS

( $p > 0.05$  en el test S-W) revelan que el efecto de interacción no fue significativo ni en comprensibilidad ( $F_{3, 72} = 0.131$ ,  $p = 0.941$ ,  $\eta_p^2 = 0.005$ ), ni en acento extranjero ( $F_{3, 72} = 0.050$ ,  $p = 0.985$ ,  $\eta_p^2 = 0.002$ ), ni en fluidez ( $F_{3, 72} = 0.039$ ,  $p = 0.990$ ,  $\eta_p^2 = 0.002$ ), lo que demuestra que el factor experiencia se comporta del mismo modo al cambiar de tramos.

Por su parte, el análisis de los efectos principales revela que las diferencias entre jueces desde esta perspectiva vuelven a confirmarse como significativas en comprensibilidad ( $F_{1, 72} = 4.779$ ,  $p = 0.032$ ,  $\eta_p^2 = 0.062$ ) y en acento extranjero ( $F_{1, 72} = 7.298$ ,  $p = 0.009$ ,  $\eta_p^2 = 0.092$ ), pero no en fluidez ( $F_{1, 72} = 0.820$ ,  $p = 0.368$ ,  $\eta_p^2 = 0.011$ ), mientras que el análisis del efecto principal de los tramos

muestra diferencias significativas en comprensibilidad ( $F_{3,72} = 16.463, p < 0.001, \eta_p^2 = 0.407$ ), en acento extranjero ( $F_{3,72} = 9.842, p < 0.001, \eta_p^2 = 0.291$ ) y en fluidez ( $F_{3,72} = 19.148, p < 0.001, \eta_p^2 = 0.444$ ). El estudio más concreto de las posibles diferencias entre los cuatro tramos de competencia en cada dimensión se realizó a continuación mediante ANOVAS de 1 factor, que detectaron, tanto en las pruebas ómnibus como en los análisis *post*

*hoc*, diferencias significativas con  $p < 0.001$  en todas las dimensiones y entre todos los tramos.

Con el fin de analizar, por tanto, cómo los 14 parámetros lingüísticos permiten diferenciar entre estos cuatro tramos de competencia percibidos por los jueces No expertos y Expertos en cada dimensión se realizan doce ANOVAS mixtos de dos factores (véase la Tabla 6). El diseño incluye, por tanto, un factor intersujetos donde

**Tabla 6:** ANOVAS mixtas de dos factores. Diferencias entre parámetros, tramos, y efecto de interacción entre parámetros y tramos.

No expertos							
Dimensiones		Parámetros de Pronunciación y disfluencias			Parámetros de Lexicogramática y discurso		
		$F^1_{(gl)}$	valor $p$	$\eta_p^2$	$F_{(gl)}$	valor $p$	$\eta_p^2$
Comprensibilidad	Parámetros	17.647 (2.351, 84.651)	.000**	.329	42.090 (3.643, 131.15)	.000**	.539
	Tramos	48.934 (3, 36)	.000**	.803	58.394 (3, 36)	.000**	.830
	Interacción	1.305 (7.054, 84.651)	.257 <sup>NS</sup>	.098	1.194 (10.93, 131.15)	.267 <sup>NS</sup>	.091
Acento	Parámetros	16.985 (2.176, 78.326)	.000**	.321	41.953 (3.757, 135.24)	.000**	.538
	Tramos	71.869 (3, 36)	.000**	.857	11.972 (3, 36)	.000**	.499
	Interacción	.806 (6.527, 78.326)	.577 <sup>NS</sup>	.063	1.151 (11.27, 135.24)	.326 <sup>NS</sup>	.088
Fluidez	Parámetros	19.938 (2.582, 92.963)	.000**	.356	44.245 (3.692, 132.90)	.000**	.551
	Tramos	40.262 (3, 36)	.000**	.770	45.497 (3, 36)	.000**	.791
	Interacción	3.033 (7.747, 92.963)	.005**	.202	1.870 (11.08, 132.90)	.048 *	.135
Expertos							
Comprensibilidad	Parámetros	18.496 (2.384, 85.832)	.000**	.339	43.568 (3.735, 134.46)	.000**	.548
	Tramos	45.322 (3, 36)	.000**	.791	34.563 (3, 36)	.000**	.742
	Interacción	1.946 (1.224, 85.832)	.071 <sup>NS</sup>	.140	1.658 (11.21, 134.46)	.088 <sup>NS</sup>	.121
Acento	Parámetros	16.955 (2.180, 78.497)	.000**	.320	41.661 (3.866, 139.16)	.000**	.536
	Tramos	39.981 (3, 36)	.000**	.769	7.250 (3, 36)	.001**	.377
	Interacción	.588 (6.541, 78.497)	.595 <sup>NS</sup>	.061	1.060 (11.60, 139.16)	.339 <sup>NS</sup>	.081
Fluidez	Parámetros	19.941 (2.523, 90.812)	.000**	.356	44.355 (3.653, 131.51)	.000**	.552
	Tramos	35.928 (3, 36)	.000**	.750	44.445 (3, 36)	.000**	.787
	Interacción	3.033 (7.568, 90.812)	.005**	.202	1.904 (10.96, 131.51)	.044 *	.137

NS = NO significativo al 5% ( $p > 0.05$ ), \* $p < 0.05$ , \*\* $p < 0.01$ .

<sup>1</sup>En todos los casos: homogeneidad de covarianzas ( $p > 0.05$  en la prueba M de Box); homogeneidad de varianzas ( $p > 0.05$  en la prueba de Levene); corrección de Greenhouse-Geisser ( $\epsilon < 0.75$  y  $p < 0.05$  en la prueba de esfericidad de Mauchly).

se comparan las notas de los hablantes por tramos de desempeño y un factor intrasujetos con medidas repetidas donde se comparan los sujetos en los 14 parámetros evaluados. Los resultados de estos ANOVAS ( $p > 0.05$  en el test S-W) muestran que existen diferencias estadísticamente significativas en todos los casos en los parámetros de «Pronunciación y disfluencias» y en los de «Lexicogramática y discurso» ( $p < 0.001$ ), con tamaños del efecto muy grandes. Esto es así también entre los tramos de desempeño de los hablantes en todas las dimensiones y en los dos grupos de evaluadores ( $p < 0.001$ ). El análisis del efecto de interacción entre parámetros y tramos ofrece diferencias significativas en la dimensión de fluidez en todos los casos ( $p < 0.05$ ) con tamaños del efecto moderados, aunque ligeramente más grandes cuando intervienen los parámetros de «Pronunciación y disfluencias». No hay un efecto de interacción en comprensibilidad y acento, lo que indica que los parámetros muestran perfiles similares en los tramos, sin que se produzcan cruces estadísticamente significativos entre ellos.

La Tabla 7 muestra a continuación los resultados del análisis de los efectos simples (con corrección de Bonferroni) que se producen según las clasificaciones que realizan los **jueces No expertos** en las tres dimensiones. Como se puede observar, en *comprensibilidad* los tres parámetros de «Ritmo», «Repeticiones, reparaciones y falsos comienzos», y «Coherencia discursiva», permiten distinguir entre todos los tramos de esta dimensión. En el lado opuesto, los segmentos solo permiten distinguir entre los tramos Inicial y Avanzado/Superior, y entre Intermedio y Superior. Además, excepto los segmentos, el resto de parámetros de pronunciación distinguen mejor entre el tramo Inicial y el resto de tramos que los parámetros de lexicogramática y discurso, donde esta distinción se produce mejor en el extremo opuesto, entre los tramos Avanzado y Superior de comprensibilidad.

Para el *acento extranjero* los parámetros de pronunciación muestran una mejor capacidad distintiva que los parámetros de lexicogramática y discurso. El «Acento», el «Ritmo» y la «Entonación» permiten distinguir entre los cuatro tramos de acento extranjero, y las «Vocales y consonantes» entre todos los tramos excepto entre el Intermedio y el Avanzado. Por su parte, los parámetros de lexicogramática y discurso solo permiten distinguir entre el tramo Inicial y Avanzado/Superior, y entre el Intermedio y el Superior.

En la *fluidez*, los cuatro parámetros de «Pausas», «Repeticiones, reparaciones y falsos comienzos», «Riqueza léxica» y «Profundidad discursiva» permiten distinguir entre todos los tramos de esta dimensión. Tras ellos se sitúa el «Tempo», que permite distinguir entre todos los tramos excepto entre Avanzado y Superior, y la «Cohesión discursiva», que permite distinguir entre todos los tramos excepto entre Inicial e Intermedio.

La Tabla 8 muestra finalmente los resultados de los mismos análisis según la clasificación de los **jueces Expertos**. En *comprensibilidad*, solo el «Ritmo» permite distinguir entre los cuatro tramos de esta dimensión. Los parámetros de «Lexicogramática y discurso» distinguen

de forma significativa todos los tramos entre sí a partir del tramo Intermedio, mientras que ninguno distingue entre Inicial e Intermedio, algo que sí hacen todos los parámetros de pronunciación, excepto los «Segmentos» y la «Entonación».

En *acento extranjero* ninguno de los 14 parámetros permite distinguir entre los cuatro tramos de esta dimensión, y en los parámetros de «Lexicogramática y discurso» solo el tramo Inicial muestra diferencias significativas con los tramos Avanzado y Superior. Los parámetros de «Pronunciación» muestran más diferencias significativas entre tramos, especialmente en los parámetros segmentales y suprasegmentales de «Acento», «Ritmo» y «Entonación».

En *fluidez*, solo el parámetro de «Repeticiones, reparaciones y falsos comienzos» permite distinguir en el grupo de Expertos entre los cuatro tramos de esta dimensión. Las «Vocales y consonantes» tienen la menor capacidad discriminante entre tramos, y solo lo hacen entre el tramo Inicial y Superior, y entre el tramo Intermedio y Superior.

## 4. DISCUSIÓN DE LOS RESULTADOS

### 4.1. Incidencia de los parámetros lingüísticos en las dimensiones

Los resultados de este estudio muestran que, cuando los oyentes evalúan la *comprensibilidad*, tienden a considerar por lo general un gran número de variables lingüísticas para acceder al significado de la forma más rápida y eficiente posible. Esto determina que la comprensibilidad sea una dimensión que se asocia, a través de valoraciones perceptivas, con variados parámetros lingüísticos, como la pronunciación, el léxico, la gramática, las marcas de disfluencia o el discurso, tal y como se ha descubierto en otras investigaciones (Isaacs y Trofimovich, 2012; Saito *et al.*, 2016, 2017). En cualquier caso, desde un punto de vista cuantitativo algunos parámetros tienen mayor incidencia que otros, y la comprensibilidad tiende a ofrecer asociaciones más altas con los aspectos lexicogramaticales y discursivos para los Expertos (especialmente con la «Corrección gramatical» y la «Precisión léxica») y con los parámetros de pronunciación para los No expertos (sobre todo con el «Ritmo» y las «Pausas»). Este comportamiento es similar al observado por Santamaría Busto (2015) con una tarea de menor complejidad –en este caso un monólogo sobre un tema familiar–, por lo que parece que estas diferencias de sensibilidad entre unos jueces y otros tienden a mantenerse independientemente de la prueba evaluada, aunque el «Ritmo» y el «Acento» para los Expertos, y especialmente el «Tempo» para los dos grupos, ganan en importancia cuando se puntúa una tarea más compleja.

El papel destacado del «Tempo» aparece recogido también en otros estudios, y se ha señalado considerando tanto la percepción que tienen los aprendientes de L2 sobre la velocidad de habla de los nativos (Derwing y Munro, 2001; Schwab, 2014), como las valoraciones que emiten los hablantes nativos sobre los tempos producidos

**Tabla 7:** Medias (desviaciones estándar) de los parámetros de «Pronunciación y disfluencias» y de «Lexicogramática y discurso» en los cuatro tramos de comprensibilidad, acento extranjero y fluidez de los hablantes, según la clasificación de los jueces No expertos.

Niveles	Comprensibilidad														
	Pronunciación					Disfluencias			Léxico			Gramática		Discurso	
	Segmentos	Acento	Ritmo	Entonac.	Tempo	Pausas	Repet. Rep. y FF. CC.	Riqueza Léxica	Precisión Léxica	Complej. Sintáctica	Corrección Gramat.	Coheren.	Cohesión	Profund.	
<i>Inic.<sup>a</sup></i>	3.22 (1.01)	3.49 (0.64)	3.23 (0.45)	3.34 (0.61)	3.33 (1.02)	3.08 (0.62)	3.70 (0.85)	3.34 (0.58)	3.03 (0.55)	2.99 (0.50)	2.70 (0.47)	3.68 (0.55)	2.97 (0.42)	3.04 (0.54)	
<i>Int.<sup>b</sup></i>	4.04 (1.12)	4.33 (0.64)	4.10 (0.89)	4.19 (0.76)	4.03 (0.95)	4.21 (0.77)	4.03 (0.87)	4.11 (0.76)	3.72 (0.85)	3.03 (0.05)	3.48 (0.84)	4.55 (0.86)	3.75 (0.79)	3.95 (0.95)	
<i>Av.<sup>c</sup></i>	4.83 (0.68)	5.17 (0.49)	5.02 (0.44)	4.99 (0.50)	5.78 (0.43)	5.26 (0.65)	6.04 (0.64)	5.46 (0.56)	4.88 (0.65)	5.01 (0.80)	4.89 (0.8?)	5.53 (0.70)	4.96 (0.84)	5.20 (0.70)	
<i>Sup.</i>	5.77 (0.99)	5.85 (0.59)	5.87 (0.47)	5.71 (0.65)	6.60 (0.52)	6.10 (0.73)	6.91 (0.37)	6.86 (0.66)	6.50 (0.07)	6.63 (0.66)	6.66 (0.67)	7.16 (0.48)	6.46 (0.56)	6.81 (0.74)	
<b>Acento extranjero</b>															
<i>Inic.<sup>a</sup></i>	2.77 (0.79)	3.37 (0.52)	3.11 (0.44)	3.21 (0.47)	3.51 (1.15)	3.27 (0.06)	3.98 (1.02)	3.64 (0.09)	3.27 (0.84)	3.31 (0.07)	3.05 (0.71)	4.05 (0.88)	3.27 (0.83)	3.38 (1.06)	
<i>Int.<sup>b</sup></i>	4.26 (0.50)	4.38 (0.45)	4.12 (0.59)	4.13 (0.36)	4.72 (1.17)	4.06 (0.93)	4.81 (1.37)	4.38 (1.40)	4.09 (1.42)	4.04 (1.51)	3.72 (1.64)	4.74 (1.57)	3.99 (1.50)	4.19 (1.64)	
<i>Av.<sup>c</sup></i>	4.73 (0.48)	5.05 (0.34)	5.00 (0.26)	5.03 (0.43)	5.77 (0.41)	5.12 (0.47)	5.91 (0.57)	5.25 (1.00)	4.72 (1.03)	4.93 (1.10)	4.72 (1.13)	5.51 (0.86)	4.86 (1.01)	5.15 (1.00)	
<i>Sup.</i>	6.11 (0.54)	6.04 (0.37)	6.00 (0.27)	5.07 (0.43)	6.54 (0.39)	6.19 (0.68)	6.79 (0.51)	6.49 (0.97)	6.04 (1.25)	6.16 (1.03)	6.23 (1.17)	6.71 (0.97)	6.01 (0.96)	6.29 (1.06)	
<b>Fluidez</b>															
<i>Inic.<sup>a</sup></i>	3.46 (0.97)	3.65 (0.74)	3.31 (0.49)	3.45 (0.69)	3.29 (0.92)	3.03 (0.57)	3.60 (0.80)	3.27 (0.55)	3.10 (0.84)	3.00 (0.59)	2.74 (0.59)	3.75 (0.67)	3.01 (0.54)	3.00 (0.56)	
<i>Int.<sup>b</sup></i>	3.98 (1.18)	4.31 (0.75)	4.07 (0.86)	4.20 (0.83)	4.73 (0.88)	4.15 (0.64)	4.87 (0.69)	4.32 (0.89)	3.82 (0.88)	3.08 (0.94)	3.65 (1.09)	4.62 (0.99)	3.82 (0.95)	4.05 (0.93)	
<i>Av.<sup>c</sup></i>	4.71 (0.98)	5.04 (0.69)	5.03 (0.67)	4.09 (0.66)	5.93 (0.30)	5.28 (0.50)	6.03 (0.44)	5.33 (0.62)	4.72 (0.75)	4.91 (0.65)	4.64 (0.82)	5.33 (0.83)	4.82 (0.74)	5.21 (0.91)	
<i>Sup.</i>	5.72 (1.10)	5.84 (0.62)	5.81 (0.58)	5.69 (0.69)	6.58 (0.53)	6.19 (0.63)	6.99 (0.37)	6.85 (0.67)	6.48 (0.87)	6.66 (0.66)	6.69 (0.67)	7.11 (0.46)	6.48 (0.57)	6.75 (0.72)	

<sup>a</sup>Las celdas con trama indican diferencias significativas al 5 % entre el tramo Inicial y otros tramos.

<sup>b</sup>El relleno de color amarillo indica diferencias significativas al 5 % entre el tramo Intermedio y los tramos Avanzado o Superior.

<sup>c</sup>El relleno de color azul indica diferencias significativas al 5 % entre los tramos Avanzado y Superior.

Comparaciones con corrección de Bonferroni.

**Tabla 8:** Medias (desviaciones estándar) de los parámetros de «Pronunciación y disfluencias» y de «Lexicogramática y discurso» en los cuatro tramos de comprensibilidad, acento extranjero y fluidez de los hablantes, según la clasificación de los jueces Expertos.

Niveles	Comprensibilidad																			
	Pronunciación				Disfluencias				Léxico				Gramática				Discurso			
	Segmentos	Acento	Ritmo	Entonac.	Tempo	Pausas	Repet. Rep. y FFCC.	Riqueza Léxica	Precisión Léxica	Complej. Sintáctica	Corrección Gramat.	Coheren.	Cohesión	Profund.						
<i>Inic.<sup>a</sup></i>	3.27 (1.06)	3.65 (0.64)	3.26 (0.48)	3.45 (0.61)	3.34 (1.02)	3.06 (0.61)	3.61 (0.80)	3.45 (0.59)	3.26 (0.78)	3.14 (0.62)	2.83 (0.60)	3.87 (0.69)	3.12 (0.57)	3.16 (0.63)						
<i>Int.<sup>b</sup></i>	3.99 (1.11)	4.16 (0.81)	4.07 (0.91)	4.07 (0.87)	4.81 (0.97)	4.22 (0.75)	4.92 (0.77)	3.99 (0.85)	3.48 (0.80)	3.68 (0.90)	3.35 (0.87)	4.36 (0.92)	3.61 (0.82)	3.83 (1.00)						
<i>Av.<sup>c</sup></i>	4.74 (0.73)	5.17 (0.49)	4.99 (0.44)	4.97 (0.51)	5.79 (0.44)	5.29 (0.66)	6.11 (0.69)	5.68 (0.77)	5.10 (0.87)	5.27 (1.07)	5.07 (1.05)	5.77 (0.85)	5.18 (1.04)	5.45 (0.98)						
<i>Sup.</i>	5.86 (0.84)	5.85 (0.59)	5.91 (0.39)	5.73 (0.61)	6.59 (0.53)	6.08 (0.75)	6.85 (0.42)	6.64 (0.86)	6.28 (1.05)	6.37 (0.84)	6.47 (0.87)	7.01 (0.66)	6.24 (0.74)	6.57 (0.91)						
<b>Acento extranjero</b>																				
<i>Inic.<sup>a</sup></i>	2.77 (0.80)	3.35 (0.42)	3.16 (0.44)	3.21 (0.44)	3.61 (1.21)	3.34 (0.87)	4.17 (0.97)	3.73 (0.85)	3.41 (0.84)	3.35 (0.85)	3.12 (0.68)	4.02 (0.91)	3.32 (0.80)	3.48 (1.03)						
<i>Int.<sup>b</sup></i>	4.27 (0.54)	4.54 (0.61)	4.27 (0.93)	4.27 (0.53)	4.85 (1.34)	4.24 (1.12)	4.79 (1.60)	4.62 (1.77)	4.33 (1.92)	4.31 (1.81)	4.10 (2.01)	5.05 (1.76)	4.23 (1.76)	4.41 (1.97)						
<i>Av.<sup>c</sup></i>	4.87 (0.60)	5.18 (0.67)	5.04 (0.59)	5.09 (0.69)	5.69 (0.75)	5.03 (0.80)	5.90 (0.79)	5.34 (1.13)	4.87 (1.24)	5.04 (1.24)	4.69 (1.47)	5.55 (1.01)	4.96 (1.14)	5.22 (1.09)						
<i>Sup.</i>	5.94 (0.76)	5.77 (0.57)	5.76 (0.46)	5.67 (0.60)	6.38 (0.64)	6.03 (0.78)	6.63 (0.62)	6.08 (1.08)	5.51 (1.22)	5.75 (1.13)	5.81 (1.15)	6.39 (1.01)	5.63 (1.08)	5.89 (1.17)						
<b>Fluidez</b>																				
<i>Inic.<sup>a</sup></i>	3.36 (0.91)	3.60 (0.70)	3.30 (0.49)	3.44 (0.67)	3.25 (0.89)	3.04 (0.58)	3.65 (0.81)	3.21 (0.49)	2.96 (0.58)	2.87 (0.40)	2.66 (0.45)	3.55 (0.47)	2.87 (0.33)	2.86 (0.41)						
<i>Int.<sup>b</sup></i>	4.21 (1.28)	4.51 (0.86)	4.16 (0.88)	4.28 (0.83)	4.78 (0.82)	4.23 (0.72)	4.97 (0.84)	4.71 (0.77)	4.27 (0.82)	4.28 (0.78)	4.03 (1.01)	5.10 (0.69)	4.17 (0.81)	4.38 (0.80)						
<i>Av.<sup>c</sup></i>	4.51 (0.96)	4.87 (0.67)	4.89 (0.74)	4.81 (0.73)	5.91 (0.29)	5.27 (0.66)	5.95 (0.79)	4.99 (0.96)	4.40 (0.99)	4.67 (1.07)	4.38 (1.16)	5.21 (1.00)	4.64 (1.05)	4.95 (0.94)						
<i>Sup.</i>	5.77 (0.99)	5.85 (0.59)	5.87 (0.47)	5.71 (0.65)	6.60 (0.52)	6.10 (0.73)	6.91 (0.37)	6.86 (0.66)	6.50 (0.87)	6.63 (0.66)	6.66 (0.67)	7.16 (0.48)	6.46 (0.56)	6.81 (0.74)						

<sup>a</sup>Las celdas con trama indican diferencias significativas al 5 % entre el tramo Inicial y otros tramos.

<sup>b</sup>El relleno de color amarillo indica diferencias significativas al 5 % entre el tramo Intermedio y los tramos Avanzado o Superior.

<sup>c</sup>El relleno de color azul indica diferencias significativas al 5 % entre los tramos Avanzado y Superior.

Comparaciones con corrección de Bonferroni.

por los hablantes de L2 (Munro y Derwing, 2001). Es posible, entonces, que las características de una tarea de cierta complejidad como es la descripción de una historia haya dirigido la atención de los oyentes hacia este parámetro en la evaluación de la comprensibilidad, bien por el uso de tempos demasiado lentos que afectan a esta dimensión, bien por el uso de tempos adecuados que la favorecen. Este resultado coincide con el alcanzado por otras investigaciones para el inglés con tareas similares a esta (Saito *et al.*, 2017; Crowther *et al.*, 2018) o más complejas (Kang *et al.*, 2010), en los que el tempo es una de las variables que más incide en la comprensibilidad de los hablantes.

Por su parte, el *acento extranjero* es una dimensión que, en línea con otros estudios (Trofimovich y Isaacs, 2012; Saito *et al.*, 2017; Crowther *et al.*, 2018), está fuertemente relacionada con las «Vocales y consonantes», seguidas de la «Entonación», el «Ritmo» y el «Acento». Esto es así independientemente de la experiencia de los jueces, aunque los No expertos muestran coeficientes más elevados que los Expertos en esta dimensión. La mayor incidencia que muestran las variables fónicas en el grupo de No expertos se mantiene también en la comprensibilidad y en la fluidez, lo que confirma de nuevo la mayor sensibilidad que tienen estos oyentes para percibir el componente fonético-fonológico del acento.

Finalmente, la *fluidez* es una dimensión que, de modo similar a la comprensibilidad, aunque generalmente con coeficientes más bajos, ofrece aquí asociaciones significativas con los aspectos fónicos, temporales, lexicogramaticales y discursivos. No obstante, los resultados muestran que en los dos grupos de jueces los parámetros de «Tempo» y «Pausas», y, en menor medida, «Repeticiones, reparaciones y falsos comienzos», destacan sobre el resto. Tales resultados coinciden con los señalados por otros estudios, en los que diversas medidas de velocidad y de pausas aparecen habitualmente como variables predictoras de la fluidez de los hablantes (Derwing *et al.*, 2004; Kormos y Dénes, 2004; Rossiter, 2009), y tienen más incidencia que las reparaciones (Bosker *et al.*, 2013).

#### 4.2. Capacidad de los parámetros lingüísticos para distinguir entre tramos de desempeño

Si se considera un análisis por tramos de desempeño en *comprensibilidad*, resulta reseñable cómo tanto para los No expertos como para los Expertos los hablantes con la peor competencia en pronunciación son también los que se diferencian de forma estadísticamente significativa en esta dimensión, mientras que los que mejor manejan los parámetros lexicogramaticales y discursivos son los que se diferencian de sus compañeros de tramos de comprensibilidad inferiores. Este resultado tiene importantes implicaciones para la enseñanza, ya que, si lo que se pretende es mejorar la comprensibilidad de un hablante situado en un tramo muy malo de desempeño, el tratamiento de la pronunciación resultará lo más efectivo para que los oyentes perciban que realmente ha mejorado. Por el contrario, cuando el hablante ha superado ya un determinado umbral

de comprensibilidad, el dominio del ritmo y de los aspectos lexicogramaticales se convierte ahora en una opción más rentable para que la mejoría del hablante le permita distinguirse de forma significativa en esta dimensión.

Por su parte, en el *acento extranjero* todos los parámetros fónicos diferencian de forma significativa a los hablantes situados en el tramo más bajo de este constructo, por lo que la enseñanza de cualquier aspecto, tanto segmental como suprasegmental, permitiría progresar de forma sustancial en esta dimensión. No obstante, cuando se supera un umbral de dominio en hablantes que tenían una baja competencia fonética se produce una diferencia muy significativa entre los jueces No expertos y Expertos. Para los primeros, el «Acento», el «Ritmo» y la «Entonación» resultan importantes para caracterizar el acento extranjero en todos sus tramos, mientras que no hay un ejemplo así en los Expertos, que, de nuevo, resultan ser menos sensibles a las variaciones de estos parámetros. A la hora de abordar en el aula esta dimensión los profesores deberían tener en cuenta, por tanto, que, aunque ellos no perciban la evolución de la competencia fonética de sus estudiantes o piensen que esta ha sido irrelevante, es muy posible que fuera del aula los interlocutores que están menos familiarizados con el español de estos hablantes sí se percaten de la mejoría, aunque luego sigan siendo más exigentes evaluándolos. En el lado opuesto, el conocimiento de los aspectos lexicogramaticales y discursivos diferencian mucho menos a los hablantes en acento extranjero, y solo lo hace con los que están situados en un tramo ya avanzado en esta dimensión.

Respecto a la *fluidez*, tanto para jueces No expertos como para Expertos el desarrollo de las «Repeticiones, reparaciones y falsos comienzos», seguido del «Tempo» y de las «Pausas», permite diferenciar de forma significativa a los hablantes en todos o casi todos los tramos, mientras que el dominio de los segmentos no resulta diferencial para distinguir en esta dimensión a unos hablantes de otros. En cambio, el «Ritmo» es el parámetro de pronunciación más efectivo cuando los hablantes han alcanzado ya un tramo intermedio. Es reseñable también cómo, en el caso de los parámetros lexicogramaticales y discursivos, los hablantes se distinguen siempre de los que tienen menos fluidez una vez que el desarrollo alcanzado en estos parámetros es elevado. Por tanto, si el objetivo del profesor es mejorar la fluidez de sus alumnos de forma efectiva, debería dar prioridad inicialmente al tempo y las vacilaciones, y, una vez que su fluidez mejora, incorporar a la enseñanza el ritmo y, posteriormente, los factores lexicogramaticales y discursivos.

#### 5. CONCLUSIONES E IMPLICACIONES DEL TRABAJO

Dada la importancia que tiene la comprensibilidad para el éxito de la interacción comunicativa, resulta decisivo que se oriente a los aprendientes de L2 sobre las variables lingüísticas –singularmente diversas– que inciden en esta dimensión. En este punto, los profesores deberían saber que los aspectos que más afectan a la comprensibilidad

de sus alumnos varían ligeramente con respecto a los que señalan los nativos que no son profesores. Estos últimos, esto es, los que no tienen experiencia en la enseñanza de segundas lenguas ni poseen un alto grado de familiaridad con el español de estos hablantes, son más sensibles a los factores fonéticos, mientras que los profesores se fijan más en la gramática. Es fundamental entonces que los docentes, pero también los creadores de materiales de segundas lenguas, lleven a las clases este punto de vista de los hablantes no expertos que, al fin y al cabo, constituyen la mayoría de los interlocutores con los que los aprendientes van a relacionarse.

Los resultados de este experimento ofrecen en cualquier caso buenas noticias para los aprendientes y los profesores de segundas lenguas. Independientemente de la experiencia de los jueces, algunos aspectos lingüísticos intervienen a la vez con fuerza en varias de las dimensiones que evalúan. Así, el ritmo es un parámetro lingüístico que tiene una gran importancia en las tres dimensiones; el tempo, las pausas y la riqueza léxica repercuten con intensidad en la comprensibilidad y en la fluidez; y las vocales y las consonantes, el acento, el ritmo y la entonación inciden especialmente en el acento extranjero. Para los anglohablantes de español como L2 el ritmo es, por tanto, el parámetro lingüístico que ofrece el mayor rendimiento de todos los analizados, ya que, tanto si el objetivo es reducir el acento extranjero como mejorar la comprensibilidad o la fluidez, los aprendientes se beneficiarán igualmente si se trabaja sobre este aspecto.

La clasificación por tramos de desempeño muestra, además, que el ritmo mantiene su importancia en todos los tramos de comprensibilidad, y que esto es así tanto para los jueces No expertos como para los Expertos. Por su parte, en la dimensión de acento extranjero cualquiera de los parámetros de pronunciación, especialmente vocales y consonantes, permite diferenciar mucho mejor a los hablantes que los parámetros lexicogramaticales y discursivos, y esto es también independiente del tramo de desempeño en que se encuentren. En la fluidez, en cambio, las repeticiones, las reparaciones y los falsos comienzos son igualmente importantes en cualquiera de los tramos de desempeño de esta dimensión.

## REFERENCIAS BIBLIOGRÁFICAS

- Allison, P. D. (1999). *Multiple Regression: A Primer*. Thousand Oaks: Pine Forge Press, Inc.
- Bergeron, A. y Trofimovich, P. (2017). Linguistic Dimensions of Accentedness and Comprehensibility: Exploring Task and Listener Effects in Second Language French. *Foreign Language Annals*, 50(3), 547–566. <https://doi.org/10.1111/flan.12285>
- Bosker, H. R., Pinget, A.-F., Quene, H., Sanders, T. y de Jong, N. H. (2013). What makes speech sound fluent? The contributions of pauses, speed and repairs. *Language Testing*, 30(2), 159–175. <https://doi.org/10.1177/0265532212455394>
- Celce-Murcia, M., Brinton, D., Goodwin, J. y Griner, B. (2010). *Teaching pronunciation. A course book and reference guide* (2.ª ed.). Cambridge: Cambridge University Press.
- Constantinescu, G. (2015). Vocation of Language for International Communication – A Prediction Tool for Future Evolutions in Global Communication. *Cross-Cultural Management Journal*, 17(2), 161–174. Disponible en: [http://seaopenresearch.eu/Journals/articles/CMJ2015\\_12\\_9.pdf](http://seaopenresearch.eu/Journals/articles/CMJ2015_12_9.pdf) [Accedido el 02/06/2020].
- Crowther, D., Trofimovich, P., Isaacs, T. y Saito, K. (2018). Linguistic Dimensions of Second Language Accentedness and Comprehensibility Vary across Speaking Tasks. *Studies in Second Language Acquisition*, 40(2), 443–457. <https://doi.org/10.1017/S027226311700016X>
- Derwing, T. M. y Munro, M. J. (2005). Second Language Accent and Pronunciation Teaching: A Research-Based Approach. *TESOL Quarterly*, 39(3), 379–397. <https://doi.org/10.2307/3588486>
- Derwing, T. M. y Munro, M. J. (2001). What Speaking Rates do Non-native Listeners Prefer? *Applied Linguistics*, 22(3), 324–337. <https://doi.org/10.1093/applin/22.3.324>
- Derwing, T. M. y Munro, M. J. (2015). *Pronunciation Fundamentals: Evidence-based Perspectives for L2 Teaching and Research*. Amsterdam: John Benjamins.
- Derwing, T. M., Rossiter, M. J., Munro, M. J. y Thomson, R. I. (2004). Second Language Fluency: Judgments on Different Tasks. *Language Learning*, 54(4), 655–679. <https://doi.org/10.1111/j.1467-9922.2004.00282.x>
- Derwing, T. M., Thomson, R. I. y Munro, M. J. (2006). English pronunciation and fluency development in Mandarin and Slavic speakers. *System*, 34(2), 183–193. <https://doi.org/10.1016/j.system.2006.01.005>
- Fleissig, D. (2009). *El análisis de los errores de la concordancia de los estudiantes universitarios anglohablantes: un estudio del aprendizaje del español como segunda lengua*. Trabajo Fin de Grado (Español con Honores). University of Michigan.
- Grüter, T., Lew-Williams, C. y Fernald, A. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research*, 28(2), 191–215. <https://doi.org/10.1177/0267658312437990>
- INE (2017). Encuesta sobre la Participación de la Población Adulta en las Actividades de Aprendizaje. *EADA*, 1–6.
- Instituto Cervantes (2020). *El español en el mundo. Anuario del Instituto Cervantes 2020*. Bala Perdida Editorial, Instituto Cervantes. Disponible en: [https://cvc.cervantes.es/lengua/anuario/anuario\\_20/default.htm](https://cvc.cervantes.es/lengua/anuario/anuario_20/default.htm) [Accedido el 26/12/2020].
- Isaacs, T. y Trofimovich, P. (2012). Deconstructing Comprehensibility: Identifying the Linguistic Influences on Listeners' L2 Comprehensibility Ratings. *Studies in Second Language Acquisition*, 34(3), 475–505. <https://doi.org/10.1017/S0272263112000150>
- Kang, O., Rubin, D. y Pickering, L. (2010). Suprasegmental Measures of Accentedness and Judgments of Language Learner Proficiency in Oral English. *The Modern Language Journal*, 94(4), 554–566. <https://doi.org/10.1111/j.1540-4781.2010.01091.x>
- Kormos, J. y Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32(2), 145–164. <https://doi.org/10.1016/j.system.2004.01.001>
- Levis, J. (2005). Changing Contexts and Shifting Paradigms in Pronunciation Teaching. *TESOL Quarterly*, 39(3), 369–377. <https://doi.org/10.2307/3588485>
- Looney, D. y Lusin, N. (2019). *Enrollments in Languages Other Than English in United States Institutions of Higher Education, Summer 2016 and Fall 2016: Final Report*. Modern Languages Association of America. Disponible en: [www.mla.org/Enrollment-Report](http://www.mla.org/Enrollment-Report) [Accedido el 14/08/2020].
- Madrid, D. (1999). Errores gramaticales en la producción escrita de los angloamericanos. En A. Romero, M. E. Fernández, E. Gómez Villalba, C. L. González, M. Moreno, C. Pérez, R. Puertas y F. Romero (Eds.), *Actas del I Congreso Internacional sobre Educación Lingüística y Literaria en el Contexto del Sistema Educativo* (pp. 609–623). Granada: Grupo Editorial Universitario.
- McBride, K. (2015). Which Features of Spanish Learners' Pronunciation Most Impact Listener Evaluations? *Hispania*, 98(1), 14–30. <https://doi.org/10.1353/hpn.2015.0001>
- Muñoz Zayas, R. (2013). El aprendizaje de lenguas extranjeras en España. *EXtoikos*, 9, 63–68.
- Munro, M. J. y Derwing, T. M. (1995a). Foreign Accent, Comprehensibility, and Intelligibility in the Speech of Second Language Learners. *Language Learning*, 45(1), 73–97. <https://doi.org/10.1111/j.1467-1770.1995.tb00963.x>
- Munro, M. J. y Derwing, T. M. (1995b). Processing Time, Accent, and Comprehensibility in the Perception of Native and Foreign-Accented Speech. *Language and Speech*, 38(3), 289–306. <https://doi.org/10.1177/002383099503800305>

- Munro, M. J. y Derwing, T. M. (2001). Modeling perceptions of the accentedness and comprehensibility of L2 speech the role of speaking rate. *Studies in Second Language Acquisition*, 23(4), 451–468. <https://doi.org/10.1017/S0272263101004016>
- Nagle, C. L. y Huensch, A. (2020). Expanding the scope of L2 intelligibility research: Intelligibility, comprehensibility, and accentedness in L2 Spanish. *Journal of Second Language Pronunciation*, 6(2), 1–23.
- O'Brien, M. G. (2014). L2 Learners' Assessments of Accentedness, Fluency, and Comprehensibility of Native and Nonnative German Speech. *Language Learning*, 64(4), 715–748. <https://doi.org/10.1111/lang.12082>
- O'Brien, M. G. (2016). Methodological Choices in Rating Speech Samples. *Studies in Second Language Acquisition*, 38(3), 587–605. <https://doi.org/10.1017/S0272263115000418>
- Ortega, L. (1999). Planning and Focus on Form in L2 Oral Performance. *Studies in Second Language Acquisition*, 21(1), 109–148. <https://doi.org/10.1017/S0272263199001047>
- Pardo, A. y San Martín, R. (2010). *Análisis de datos en ciencias sociales y de la salud II*. Madrid: Síntesis.
- Pinget, A. F., Rutger Bosker, H., Quené, H. y de Jong, N. H. (2014). Native speakers' perceptions of fluency and accent in L2 speech. *Language Testing*, 31(3), 349–365. <https://doi.org/10.1177/0265532214526177>
- Piske, T., MacKay, I. y Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, 29(2), 191–215. <https://doi.org/10.006/jpho.2001.0134>
- Préfontaine, Y. y Kormos, J. (2016). A qualitative analysis of perceptions of fluency in second language French. *International Review of Applied Linguistics in Language Teaching*, 54(2), 151–169. <https://doi.org/10.1515/iral-2016-9995>
- Rossiter, M. J. (2009). Perceptions of L2 Fluency by Native and Non-Native Speakers of English. *The Canadian Modern Language Review/La Revue Canadienne des Langues Vivantes*, 65(3), 395–412. <https://doi.org/10.3138/cmlr.65.3.395>
- Saito, K. (2019). To What Extent Does Long-Term Foreign Language Education Help Improve Spoken Second Language Lexical Proficiency? *TESOL Quarterly*, 53(1), 1–26. <https://doi.org/10.1002/tesq.468>
- Saito, K. y Akiyama, Y. (2017). Linguistic Correlates of Comprehensibility in Second Language Japanese Speech. *Journal of Second Language Pronunciation*, 3(2), 199–217. <https://doi.org/10.1075/jslp.3.2.02sai>
- Saito, K., Trofimovich, P. y Isaacs, T. (2016). Second Language Speech Production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. *Applied Psycholinguistics*, 37(2), 217–240. <https://doi.org/10.1017/S0142716414000502>
- Saito, K., Trofimovich, P. y Isaacs, T. (2017). Using Listener Judgments to Investigate Linguistic Influences on L2 Comprehensibility and Accentedness: A Validation and Generalization Study. *Applied Linguistics*, 38(4), 439–462. <https://doi.org/10.1093/applin/amv047>
- Sánchez Avendaño, C. (2002). La percepción de la fluidez en español como segunda lengua. *Revista de Filología y Lingüística de La Universidad de Costa Rica*, 28(1), 137–163. <https://doi.org/10.15517/rfl.v28i1.4507>
- Santamaría Busto, E. (2015). Percepción y evaluación de la pronunciación del español como L2. *Revista Española de Lingüística*, 45(1), 175–207. <https://doi.org/10.31810/RSEL.45.1>
- Schairer, K. E. (1992). Native Speaker Reaction to Non Native Speech. *The Modern Language Journal*, 76(3), 309–319. <https://doi.org/10.2307/330161>
- Schwab, S. (2014). La perception du débit en espagnol L1 et L2: effet de la langue maternelle, du degré d'accent perçu et des compétences en L2. *Nouveaux Cahiers de Linguistique Française*, 31, 117–128. <https://doi.org/10.5167/uzh-126025>
- Szabo, M. (2006). "I Meant to Say That": How Adult Language Learners Construct Positive Identities Through Nonstandard Language Use. *TESL Canada Journal*, 24(1), 21–40. <https://doi.org/10.18806/tesl.v24i1.26>
- Tabachnick, B. G. y Fidell, L. S. (2007). *Using Multivariate Statistics* (5.ª ed.). Boston: Pearson.
- Trofimovich, P. y Isaacs, T. (2012). Disentangling accent from comprehensibility. *Bilingualism: Language and Cognition*, 15(4), 905–916. <https://doi.org/10.1017/S1366728912000168>
- Winke, P., Gass, S. y Myford, C. (2013). Raters' L2 background as a potential source of bias in rating oral performance. *Language Testing*, 30(2), 231–252. <https://doi.org/10.1177/0265532212456968>