

Revista Española de Lingüística

Órgano de la Sociedad Española de Lingüística

RSEL

45|1

Enero-Junio
2015

Edita
SeL

REVISTA ESPAÑOLA DE LINGÜÍSTICA
(RSEL)
45/1

Edita

SeL

REVISTA ESPAÑOLA DE LINGÜÍSTICA (RSEL)

ISSN: 0210-1874 · eISSN: 2254-8769

Depósito Legal: M-24.769-1971

DIRECTOR DE HONOR: D. Francisco Rodríguez Adrados (RAE y RAH).

DIRECTOR: Juan Antonio Álvarez-Pedrosa Núñez (UCM).

SECRETARIO: Luis Unceta Gómez (UAM).

CONSEJO DE REDACCIÓN: Montserrat Benítez (CSIC), José Antonio Berenguer (CSIC), Joaquín Garrido (UCM), Juana Gil Fernández (CSIC), Salvador Gutiérrez Ordóñez (U. León y RAE), Antonio Hidalgo (U. Valencia), Patricia Infante (CSIC), Manuel Leonetti (U. Alcalá), Eugenio Luján (UCM), Victoria Marrero (UNED), Ventura Salazar (U. Jaén), Esperanza Torrego (UAM).

CONSEJO ASESOR: Alberto Bernabé (UCM), Margarita Cantarero (SEL), Ramón Cerdá (UB), Victoria Escandell (UNED), Marina Fernández Lagunilla (UAM), José Manuel González Calvo (U. Extremadura), Emma Martinell (UB), Juan Carlos Moreno Cabrera (UAM), Gregorio Salvador (RAE), José Carlos de Torres (SEL), Jesús de la Villa (UAM).

A partir del número 38 (2008) la Revista Española de Lingüística ha recuperado el formato de dos fascículos al año, con periodicidad semestral. Los trabajos enviados para su publicación han de dirigirse al Secretario de la revista. Deberán ser originales e inéditos y ajustarse a las normas que aparecen en el número 38/2, así como en la página web de la Sociedad Española de Lingüística. Todos los trabajos son sometidos al dictamen de al menos dos evaluadores designados por el Consejo de Redacción, mediante informes de carácter confidencial.

Los derechos de publicación y difusión, bajo cualquier forma, son propiedad de RSEL. Todo texto publicado en la revista obliga a sus autores a no cederlo a terceros, sin autorización previa de la revista, quien sí queda autorizada a comercializarlo, debiendo entregar, en este caso, el 50% de los beneficios obtenidos a sus autores.

REDACCIÓN: Sociedad Española de Lingüística, Centro de Ciencias Humanas y Sociales del CSIC, c/ Albasanz, 26 - 28, 28037 Madrid. Correo electrónico: secretarioRSEL@gmail.com. <<http://www.sel.edu.es/>>

DISEÑO Y COMPOSICIÓN: Carmen Chíncoa & Carlos Curiá (produccionRSEL@gmail.com)

SERVICIOS DE INFORMACIÓN: Los contenidos de la RSEL son recogidos sistemáticamente en *Bibliographie Linguistique / Linguistic Bibliography*, *CINDOC – Base de datos Sumarios ISOC, CSA – Linguistic and Language Behavior Abstracts, Dialnet, Francis, Modern Language Association (MLA) Bibliography*.

ÍNDICE 45/1 (2015)

INTRODUCCIÓN

Nuevas aportaciones al estudio de la percepción del habla7
JUANA GIL - EUGENIA SAN SEGUNDO

ARTÍCULOS

Implicaciones perceptivas de la variación: la fricativa labiodental25
BEATRIZ BLECUA FALGUERAS - ASSUMPCIÓ ROST BAGUDANCH

La percepción de la cualidad de voz y los estereotipos vocales45
MARIANELA FERNÁNDEZ TRINIDAD

La prosodia audiovisual de la ironía verbal: un estudio de caso73
SANTIAGO GONZÁLEZ FUENTE

¿Son distintos el creak y la voz creaky?: Estudio perceptivo preliminar105
PATRICIA INFANTE

La percepción del habla en ruido: un reto para la lingüística129
y para la evaluación audiológica (estudio experimental)
VICTORIA MARRERO-AGUIAR

Percepción de los enunciados interrogativos de duda y de enfado153
sin apoyo visual en alumnos chinos de ELE
SÍLVIA PLANAS-MORALES

Percepción y evaluación de la pronunciación del español como L2175
ENRIQUE SANTAMARÍA BUSTO

ARTÍCULOS

PERCEPCIÓN Y EVALUACIÓN DE LA PRONUNCIACIÓN DEL ESPAÑOL COMO L2

ENRIQUE SANTAMARÍA BUSTO
New York University in Madrid

RESUMEN

Este estudio se propone analizar y comparar los principales constructos (inteligibilidad, comprensibilidad, acento extranjero, fluidez e irritabilidad) que intervienen en la percepción y evaluación de la pronunciación de una L2. Aunque en la bibliografía pueden encontrarse investigaciones que recogen las relaciones que se establecen entre algunas de estas dimensiones, aún permanece inexplorado el análisis conjunto de todas ellas, y son muy pocos los estudios que han considerado para el español como L2 una dimensión que no fuera el acento extranjero. Con el fin de analizar el papel y la relación que guardan entre sí estas dimensiones, se examinó en relación con cada una de ellas el proceso de evaluación que 44 evaluadores nativos (15 expertos y 29 no expertos) realizaban sobre ocho muestras orales de español como L2. Los resultados reflejan, entre otros datos, que los jueces tienden a sobrevalorar su capacidad para entender las muestras, que la inteligibilidad y la comprensibilidad guardan entre sí una estrecha relación, y que el acento extranjero es el constructo que más marca la diferencia en las valoraciones perceptivas entre jueces expertos y no expertos.

Palabras clave: percepción; evaluación; pronunciación de una L2; inteligibilidad; comprensibilidad; acento extranjero; fluidez; irritabilidad.

ABSTRACT

This study aims to analyze and relate the major constructs (intelligibility, comprehensibility, accentedness, fluency and irritability) operating in the perception and assessment of L2 pronunciation. Although relations between some of these dimensions have been discussed in the literature, it still remains unexplored an overall analysis of all of them. In order to evaluate the role and connections of these dimensions, the oral productions of 8 native English speakers of Spanish were examined for each of the 5 constructs by 29 novice raters and 15 experienced teachers (N=44). Results showed that intelligibility and comprehensibility are strongly correlated, that raters tend to overestimate their ability to understand the speech samples, or that accentedness makes the difference on perceptual evaluations between expert and non-expert raters.

Keywords: perception, assessment; L2 pronunciation; intelligibility; comprehensibility; accentedness; fluency; irritability.

RECIBIDO: 04/03/2014

APROBADO: 07/01/2015

1. INTRODUCCIÓN

El despertar que, en los últimos años, han experimentado los campos de la pronunciación y de la evaluación en la enseñanza de segundas lenguas ha traído aparejado un marcado interés por el estudio de los grandes constructos y variables que operan en el proceso de producción y percepción oral de una L2. A la luz de estas investigaciones, términos como inteligibilidad, comprensibilidad, acento extranjero, fluidez e irritabilidad han resurgido como elementos de análisis que, con sus variadas interdependencias, se sitúan en un triple plano que cubre la interacción de hablantes/oyentes, la reacción de los interlocutores, y la competencia oral en una lengua.

Las investigaciones realizadas hasta la fecha permiten comprobar también algunas discrepancias, como la falta de consenso sobre la definición de estos constructos, o la diversidad de metodologías empleadas para medirlos, lo que dificulta la tarea de comparar los resultados. Es posible entonces encontrar en la bibliografía diferencias conceptuales entre estudios que aparentemente remiten al mismo constructo (por ejemplo, inteligibilidad) cuando en la práctica están midiendo dimensiones diferentes (normalmente comprensibilidad). Esta incoherencia puede observarse también en el campo de la evaluación de la pronunciación, en cuyo marco se diseñan a veces descriptores y escalas en los que confluyen la inteligibilidad, la comprensibilidad y el acento extranjero, sin tener en cuenta que son dimensiones parcialmente independientes y que, por tanto, no guardan siempre entre ellas una relación directa (Derwing y Munro 2009).

Considerados los aspectos anteriores, este estudio propone analizar estos constructos, estudiar las relaciones que se mantienen entre ellos, y examinar los elementos lingüísticos que señalan los oyentes en los análisis perceptivos, a partir del uso de escalas y evaluaciones analíticas. Con ese objetivo, se diseñó un experimento que consistió en la obtención de ocho muestras orales de habla semiespontánea, de aproximadamente un minuto, producidas por anglohablantes que tienen el español como L2, así como la fabricación de dos cuestionarios que se repartieron a 44 evaluadores nativos divididos en dos grupos (15 expertos y 29 no expertos). Ambos grupos tuvieron que completar los siguientes apartados: a) transcripción de 24 frases producidas por los ocho hablantes (tres frases por hablante) extraídas de la parte inicial, media y final del minuto de muestra, con el fin de medir la inteligibilidad; y b) evaluación a través de escalas Likert de nueve niveles de la comprensibilidad, acento, fluidez e irritabilidad, unida a estudios analíticos que pretendían determinar los aspectos particulares del constructo observado por los jueces.

La hipótesis de este estudio es que se encontrarían diferencias intergrupales que afectarían al grado de inteligibilidad, a la valoración de los constructos o al distinto peso otorgado a los elementos analizados en ellos. Por otro lado, no se esperaban diferencias significativas en cuanto

a las relaciones establecidas entre los constructos dentro de cada grupo, con la consideración de que las notas de comprensibilidad estarían más correlacionadas con las notas de inteligibilidad (medidas a través de transcripciones) que con las de acento extranjero, tal como aparece de forma recurrente en la bibliografía.

2. EL ACENTO

El acento, entendido como el efecto auditivo global que produce un conjunto de rasgos fónicos presentes en el mensaje oral, es uno de los aspectos más prominentes del habla. Permite identificar no solo la procedencia social y regional de una persona, sino también las variaciones fónicas resultantes de la influencia de la L1 en una L2. En el campo de las segundas lenguas, la competencia fonética se configura por tanto como uno de los aspectos de mayor relevancia en los procesos de comunicación, pues su falta de dominio puede llevar consigo repercusiones sociales, psicológicas y comunicativas (Munro 2003).

A la luz de estos aspectos, el estudio sobre el acento extranjero¹ y la didáctica de la pronunciación están experimentando una fuerte recuperación que parece responder tanto a los intereses de profesores y alumnos, como al beneficio social y la rentabilidad económica que se derivan de ellos. En numerosos países (entre ellos España a través de los exámenes DELE) es obligatorio superar con éxito pruebas (IELTS, TOEFL, TSE, ACTFL, CAE/CPE, DELF/DALF, DSH, CELPE-Bras) que miden el dominio de la competencia oral y, dentro de ella, de la pronunciación, para poder acceder a puestos de trabajo, permisos de residencia, obtención de nacionalidad, o ingreso en la universidad.

Al mismo tiempo, la llegada del enfoque comunicativo, con su concepción de la lengua entendida como instrumento de comunicación y no como fin en sí misma, supuso una reorientación de los objetivos de enseñanza que afectó también a la pronunciación. Esto significó un cambio gradual en el foco que desde la enseñanza de la pronunciación y de la corrección fonética se había puesto tradicionalmente en la producción correcta de los sonidos y no en la transmisión comprensible del mensaje. Se marcaron así dos polos de contraste, todavía presentes en el aula, como

1. Aprovecho para señalar la diferencia entre acento y acento extranjero con los siguientes matices que aparecen en este estudio: el acento, como se ha dicho, es el efecto perceptivo global que produce el conjunto de rasgos fónicos presentes en el proceso de producción del habla. Considerado así, puede decirse que todo el mundo habla con acento, aunque este no informe necesariamente sobre el origen geográfico de una persona. El acento extranjero, por su parte, se define como el grado de diferencia fonética que presenta la emisión de un hablante con respecto a las normas nativas (Riney 2005). El acento extranjero, pero también la inteligibilidad y la comprensibilidad, de los que se hablará en el presente trabajo, son por tanto dimensiones que hacen referencia a maneras de medir y valorar ese acento.

son el principio de inteligibilidad y la búsqueda del acento nativo (Levis 2005). El enfoque comunicativo significó entonces un rescate, pero sobre todo una reconsideración de la jerarquía que ocupaban estos elementos en la producción y percepción del habla. Se empezaba a valorar que los hablantes de segundas lenguas deberían ser en, primer lugar, entendidos por sus interlocutores, y que un acento extranjero no impedía necesariamente una comunicación efectiva (Munro y Derwing 1999). Esta suerte de movimiento pendular hacia otras unidades de análisis (inteligibilidad, comprensibilidad, fluidez) hizo que se recuperara la investigación sobre ellas, y estas empezaran a aplicarse en la enseñanza y evaluación de segundas lenguas.

2.1. *La inteligibilidad*

Actualmente existe un consenso mayoritario entre investigadores y profesores de LE en aceptar un orden de importancia que empieza en la inteligibilidad y continúa con la comprensibilidad (Kennedy y Trofimovich 2008). Sin embargo la inteligibilidad ha sido tradicionalmente una dimensión «escurridiza» sobre la que no ha existido acuerdo, ni en su definición ni en la manera de medirla (Isaacs 2008). Incluso es posible comprobar en la bibliografía cómo la inteligibilidad y la comprensibilidad han sido confundidas e interpretadas de la misma forma, y lo que para unos investigadores es inteligibilidad para otros es comprensibilidad (Gass y Varonis 1984; Ludwig 1982).

En su deseo de definir el objeto de estudio, Smith y Nelson 1985 sugirieron una conceptualización de estas dos dimensiones que diez años después sería retomada por Munro y Derwing 1995a y 1995b y Derwing y Munro 1997, constituyéndose un marco operativo que a partir de entonces sería el que tendría una mayor implantación. De acuerdo con este enfoque, que es el que se ha seguido en este trabajo, la inteligibilidad se define como la cantidad de muestra oral que el oyente es capaz de entender y suele medirse examinando la proporción de palabras que es capaz de transcribir correctamente, aunque se han usado también otros métodos². El campo de acción está, por tanto, en el reconocimiento de palabras y estructuras, y no en la construcción y el procesamiento de sus significados en un contexto, lo que sería propio de la comprensibilidad, como después veremos.

2. Por ejemplo: mediciones holísticas, hechas a través de escalas Likert (Fayer y Krasinski 1987); colocando una marca sobre escalas continuas (Isaacs 2008); a través de ejercicios de huecos (Chen 2011); respondiendo a preguntas de comprensión (Hahn 2004); mediante el uso de tecnologías de reconocimiento del habla (Raux y Kawahara 2002); deteniendo un vídeo cada vez que algo impedía la comprensión (Gallego 1990); o contestando a preguntas de «verdadero» o «falso» (Munro y Derwing 1995b; Kennedy y Trofimovich 2008).

Dada la importancia de la inteligibilidad como dimensión que permite o impide la comunicación, uno de los principales objetivos de los investigadores ha sido aislar y conocer los aspectos lingüísticos que la imposibilitan. Para el inglés, que es la lengua dominante en este tipo de estudios, se han señalado, entre otros, y desde diversas LI, la posición acentual (Hahn 2004); la duración de los segmentos, las pausas y la F0 (Powell Crocker 2010); o la duración silábica y la F0 (Winters y O'Brien 2013). En cambio, el número de estudios relacionados con la inteligibilidad del español como L2 es muy limitado y está más relacionado con la comprensibilidad (Schairer 1992) y el análisis de la producción oral de hablantes con acento extranjero. Este último tipo de estudios, más numeroso, no determinan por lo general las dimensiones concretas a las que puede aplicarse su análisis, es decir, si los errores de pronunciación que se producen impiden la comunicación, la dificultan, o no la dificultan pero no se corresponden con una producción nativa (Llisterri 2003). Por ello, como señala este autor, los estudios sobre estas dimensiones (realizados sobre cada particular L1-L2) se antojan muy necesarios como medios de análisis que permitirían «priorizar contenidos» y «proporcionar al examinador una orientación sobre los errores más probables en los que debe centrar su atención» (Llisterri 2003, p. 7).

Por lo general, la mayoría de las investigaciones han señalado que los aspectos suprasegmentales tienen un papel más importante que los aspectos segmentales en la percepción que afecta a cualquiera de las dimensiones del acento (Field 2005). Eso sí, es preciso señalar que los estudios que han analizado las correlaciones entre los «errores» fónicos y las notas que hablantes nativos ponían en cada una de estas dimensiones lo han hecho casi siempre con hablantes nativos en el papel de jueces evaluadores, dejando a un lado a los no nativos (Jenkins 2000). Estudios que analicen por tanto este tipo de relaciones son también muy necesarios, pues, como señala esta última autora (Jenkins 2000, pp. 80-84), los evaluadores no nativos suelen usar estrategias de procesamiento diferentes a los nativos (en su caso de abajo-arriba «bottom-up approach», y no de arriba-abajo «top-down approach» tal y como realizarían los hablantes nativos). A ello habría que añadir que los hablantes no nativos muestran más dificultades en compensar los errores de pronunciación usando la información sintáctica y contextual presente en la emisión fónica (Jenkins 2000). Es por ello probable que futuros estudios realizados en esta línea confirmen de nuevo estos resultados y que, para este grupo, sean los aspectos segmentales los que causen más problemas de comprensión.

Por otro lado, aunque la bibliografía ha relacionado tradicionalmente niveles altos de acento extranjero con una baja inteligibilidad, Derwing y Munro 1997 demostraron que era posible tener un fuerte acento extranjero y a la vez entender perfectamente lo que el hablante decía. Lo contrario, sin embargo, no suele pasar, de modo que a niveles bajos de inteligibilidad les corresponde casi siempre un fuerte acento extranjero. Por tanto, es

preciso tener en cuenta que estas dimensiones mantienen siempre relaciones entre ellas, pero son parcialmente independientes. Estos resultados tienen de este modo serias implicaciones para la validación del diseño de escalas y descriptores de evaluación en los que todavía se establece una relación directa que no siempre se produce.

2.2. *La comprensibilidad*

La comprensibilidad se ha definido como la percepción sobre el grado de dificultad que tiene una muestra para ser entendida (Derwing 2008, p. 185). Si la inteligibilidad se mide a través de transcripciones u otros medios de evaluación, la comprensibilidad se hace, al igual que en el acento extranjero, la fluidez o la irritabilidad, a través de juicios emitidos por los jueces evaluadores mediante el uso de escalas de niveles. Las escalas de este tipo, llamadas escalas Likert (Busch 1993), emplean una graduación numérica cuyos extremos están indicados con números y descriptores que presentan una oposición informativa (p. ej. sin acento extranjero y acento extranjero extremadamente marcado), aunque el orden en que se presenta esta información, así como el número de niveles de la escala, puede variar de unos autores a otros.

La percepción de la comprensibilidad corresponde por tanto con el grado de esfuerzo y de tiempo que invierte el interlocutor para procesar el contenido del mensaje, aunque este sea completamente inteligible (Munro y Derwing 1995b). Niveles bajos de comprensibilidad supondrían así la conciencia de que el procesamiento y la comprensión de la información requieren un tiempo extra y un aumento de la concentración. Teniendo en cuenta que la comprensibilidad mide el grado de dificultad, los investigadores observaron que los interlocutores no reaccionaban y evaluaban exclusivamente a partir de los aspectos fónicos, sino que también lo hacían a partir del grado de fluidez y de corrección gramatical, léxica y discursiva, por lo que en los estudios que analizan esta dimensión (Isaacs y Trofimovich 2012) ha sido habitual considerar estos elementos. Estos aspectos también han sido considerados como factores que impiden la inteligibilidad (Derwing y Munro 1997) y aumentan el grado de acento extranjero (Trofimovich e Isaacs 2012), aunque parecen operar con más fuerza en la comprensibilidad.

Dado el carácter central de la comprensibilidad como agente que facilita la comunicación, el conocimiento de los aspectos que más inciden de forma negativa en ella ha sido el objetivo de un número cada vez más creciente de estudios sobre el tema. Isaacs y Trofimovich 2012, en un estudio en el que se correlacionaron las notas de evaluadores nativos expertos y no expertos con las características lingüísticas de un grupo de hablantes que tenía el inglés como L2, encontraron que los dos grupos de evaluadores identificaron de forma unánime los errores léxicos y gramaticales como

causa principal que impedía la comprensibilidad, mientras que los elementos segmentales, la posición acentual, la realización silábica y el ritmo eran errores que se identificaban con el acento extranjero³. En esta misma línea, otros estudios habían avanzado ya la importancia que los dominios gramatical y léxico parecen tener sobre la comprensibilidad (Derwing y Munro 1997; Munro y Derwing 1999), lo que demuestra, como señalábamos, que la mejora de la comprensibilidad no depende exclusivamente del dominio de la competencia fonética.

Estos resultados parecen indicar, pues, la necesidad de integrar la enseñanza de la pronunciación y la corrección fonética en los demás contenidos lingüísticos analizados habitualmente en este dominio, en línea con los planteamientos metodológicos más actuales (Santamaría Busto 2013), así como el deber de conocer mejor la relación de estos elementos a la hora de diseñar escalas y descriptores de evaluación (Llisterri 2003).

Los aspectos fónicos, especialmente los suprasegmentales, tienen también una fuerte correlación con la comprensibilidad y han sido los más estudiados, especialmente en el campo del inglés como L2 (Kang 2010). En el caso del español, es de destacar el estudio realizado por Schairer 1992 sobre anglohablantes de Estados Unidos que tienen el español como LE. La autora señala en primer lugar la producción de los segmentos vocálicos y posteriormente la ausencia de enlaces (resilabificaciones producidas entre las vocales en inicio de palabra y las consonantes finales de las palabras que las preceden), como las características fónicas de la interlengua de los anglohablantes que causan mayores problemas de comprensión entre los hablantes nativos. Considerando estos estudios y la importancia de cada una de estas dimensiones, se hace muy necesario seguir trabajando en el conocimiento de los aspectos lingüísticos que las definen.

2.3. *El acento extranjero*

Como ya se dijo, el acento extranjero se define como el grado de diferencia fonética que presenta la emisión de un hablante con respecto a las normas nativas (Riney 2005). Comparada con las dos dimensiones anteriores, Derwing y Munro 2009, p. 480, señalan que el acento extranjero debe entenderse entonces como una cuestión de diferencia, la comprensibilidad como un grado de esfuerzo y la inteligibilidad como el resultado final.

El acento extranjero es también la dimensión que guarda una relación menos directa con la inteligibilidad y la comprensibilidad, y el grado en el que se manifiesta no afecta necesariamente a las otras dos, que sí suelen mantener una relación más directa. Debido a su prominencia, ha sido

3. En este, como en todos los estudios, es preciso tener siempre en cuenta la L1 del hablante, en este caso el francés.

también la dimensión más estudiada en la bibliografía sobre segundas lenguas, bien en lo que concierne a la reacción de los interlocutores, bien en lo que se refiere al proceso de producción y percepción de una L2, a su enseñanza y corrección, o a las características lingüísticas (principalmente fónicas) que caracterizan la interlengua en esta dimensión.

El acento extranjero no impide necesariamente la inteligibilidad ni la comprensibilidad, aunque por sí mismo puede tener tantas repercusiones sociales en los procesos de comunicación que la mejora en esta dimensión suele alcanzar un elevado grado de consenso entre los hablantes de segundas lenguas (Derwing 2003). Una consecuencia de ello es que la enseñanza y la corrección fonética de los aspectos que inciden en esta dimensión suele verse por profesores y alumnos como algo positivo (Barrera Pardo 2004). Hablar con acento extranjero puede tener beneficios, ya que permite saber al interlocutor que el hablante carece de una competencia fonética nativa y puede así adaptar su discurso (Gass y Varonis 1984). Sin embargo también puede tener repercusiones negativas.

En una revisión de experiencias y episodios recopilados por Lippi-Green 2004, la autora destaca los costes que tuvieron que pagar en Estados Unidos no solo algunos hablantes que tenían el inglés como L2 y lo hablaban con un determinado acento extranjero, sino incluso hablantes que tenían el inglés como L1 pero procedían de variedades dialectales distintas a las del entorno en las que se desarrollaban sus relaciones sociolaborales. Son casos en los que operan prejuicios y estereotipos sociolingüísticos que pueden llegar a la discriminación por razón del acento (Munro 2003).

Estas actitudes del interlocutor, tanto de los hablantes nativos como incluso de los no nativos (Fayer y Krasinski 1987), han sido muy estudiadas por la sociolingüística. La pronunciación de una persona forma parte de su identidad y escuchar un acento extranjero implica conocer información sobre esa persona. Como señalan Derwing y Munro 2009, p. 486, el simple hecho de que un oyente perciba un acento extranjero y sepa que su interlocutor no es nativo puede producir niveles más bajos de comprensibilidad e inteligibilidad, por la sencilla razón de que «ha decidido» que no va a entenderlo.

Esta situación ha permitido en los últimos años sacar a la luz la relevancia que tiene también el papel del interlocutor como interventor necesario en el éxito del mensaje (Rajadurai 2007). El oyente, por tanto, tiene asimismo un papel activo y el éxito de la comunicación depende también de su voluntad para comprometerse en la interacción. Esto permite en cierta medida situar el foco no exclusivamente en el hablante (como tradicionalmente suele hacerse en los estudios de evaluación) y dirigirlo un poco más hacia el oyente (Derwing, Rossiter y Munro 2002).

Por otra parte, es preciso tener en cuenta que la inmensa mayoría de los hablantes que aprenden una segunda lengua a partir de una determinada edad, poseen un acento extranjero (Singleton y Ryan 2004), ya que en sí

mismo es un fenómeno complejo sujeto a múltiples variables (Gil 2007)⁴. Esta dificultad para perder el acento extranjero trae consigo importantes implicaciones para el campo de la enseñanza de la pronunciación y de la corrección fonética que es preciso considerar. Por un lado, permiten establecer en el aula objetivos realistas y, por otro, subrayan la necesidad de proporcionar conocimientos y herramientas que ayuden al profesor a saber dónde centrar su atención, ya que podemos correr el riesgo de dirigirla a aspectos que resultan llamativos (por ejemplo, la pronunciación de la vibrante española), que sin embargo pueden no ser tan decisivos para la inteligibilidad o la comprensibilidad.

Finalmente, es preciso considerar también que los ejemplos de interferencia fonética pueden ser a su vez múltiples y pluriparamétricos. Solo en la entonación del español podemos encontrar, entre otros muchos: ausencia de inflexiones tonales no conclusivas en declarativas no enfáticas, como sería esperable en la mayor parte –no todas– de las variedades dialectales del español, por lo que el oyente interpreta que aún no ha concluido el turno de palabra; presencia de múltiples picos tonales en hablantes de italiano (coincidentes con sílabas tónicas) en la zona media del contorno tonal, y tonema descendente a partir de la pretónica y no de la tónica, como suele ocurrir en español; tendencia de los sinohablantes a convertir cada palabra en un grupo fónico o a transmitir significados expresivos muy poco marcados; marcación muy prominente de los focos en anglohablantes de español como L2, etc.

La forma precisa en que se agrupan los sonidos, y la sucesión de los diferentes contornos tonales (con sus acentos, rangos, movimientos e inflexiones) son las que originan la melodía de la lengua y las principales responsables, junto con el ritmo, de lo que llamamos acento extranjero (Cantero 2003).

La cuestión es, sin embargo, determinar cuáles de estos aspectos (u otros) inciden en primer lugar en la inteligibilidad, para elaborar pautas de actuación didáctica en el aula. Una vez alcanzada esta base mínima se avanzaría hacia los errores que dificultan esa comunicación, para pasar finalmente al estudio de las peculiaridades segmentales y suprasegmentales que ayudan a mejorar la calidad global de la emisión fónica, y con ella a alcanzar el dominio de la competencia fonética.

4. Algunas de ellas, por citar a las más importantes, son: la edad, la aptitud, la motivación, la actitud hacia la lengua meta, la identificación por la L1, los factores afectivos y psicosociales, la calidad y cantidad de *input*, los tipos de instrucción, el contexto y los estilos de aprendizaje, o la distancia tipológica entre las L1 y L2 (véase también Piske y otros 2001).

2.4. *La fluidez*

La fluidez es «la habilidad de procesar la lengua con soltura y coherencia, sin excesivas pausas ni titubeos» (Instituto Cervantes 2008) y a una velocidad constante y similar a la que se esperaría normalmente de un nativo en ese mismo contexto. La conjunción de todas estas variables ha hecho que la fluidez se haya considerado una dimensión compleja en la que suelen investigarse factores lingüísticos, temporales y comunicativos sobre los que no existe siempre consenso ni en su definición ni en la manera de medirlos (Kormos y Dénes 2004). Esta definición del término ha traído consigo la idea de que hablar «con soltura», «sin pausas» y a la «velocidad de un nativo», encajaba perfectamente con el dominio de la competencia oral de una lengua, pues en niveles de baja competencia los problemas léxicos, gramaticales o discursivos llevaban aparejados ciertas vacilaciones, reformulaciones, pausas, repeticiones y falsos comienzos que afectaban al tiempo y al modo de transmisión del mensaje. Esto es lo que llevó a Lennon 1990 a afirmar que en el campo de las segundas lenguas existen dos tipos de fluidez: la que se asocia en términos globales con un elevado dominio de competencia oral, muy bien considerada; y la que se define en términos más específicos, la cual engloba las sensaciones perceptivas con las que se manifiesta.

Estas sensaciones pueden ser de varios tipos, aunque las más estudiadas dentro de esta dimensión han sido las producidas por las marcas temporales. Tales marcas engloban no solamente el uso de las pausas (posición, duración, realización y número de pausas), sino también otros aspectos relacionados con la producción del habla por unidad de tiempo, bien cuando se consideran las pausas vacías (tempo o velocidad de elocución), bien cuando se prescindir de ellas (velocidad de articulación). El uso de la pausa sirve también para estudiar otras unidades de medida que pueden dar información sobre la fluidez, tales como la duración media de los grupos fónicos (aunque no existe consenso en cuanto a la duración de las pausas que los delimitan), o ciertas unidades indicativas de disfluencia, como pueden ser las repeticiones, las reparaciones, o los falsos comienzos.

Los aspectos fónicos o, para ser más exactos, algunos aspectos fónicos de la producción del habla intervienen también por sus cualidades en la definición de la fluidez. Los más habituales son los enlaces o las resilabificaciones, mediante los cuales las vocales sucesivas o las vocales precedidas de consonantes entre palabras diferentes se reestructuran silábicamente por medio del enlace de unas con otras; o la entonación, que con sus cambios y movimientos tonales agrupa los sonidos del habla en bloques (y no en palabras o sílabas aisladas).

Uno de los principales objetivos de la bibliografía que estudia la intervención de todas estas variables en el grado de fluidez en una segunda lengua ha sido el tratar de medirlas y cuantificarlas con el fin de encontrar

los aspectos que más inciden en la falta de dominio de esta dimensión. Por otra parte, también se han analizado variables no lingüísticas, como el tipo de tarea o el contexto de comunicación. Por ejemplo, parece ser que los hablantes tienen la sensación de hablar con más fluidez en las interacciones con nativos (Ejzemberg 2000, *apud* Derwing y Rossiter 2004), y que la carga cognitiva, el léxico, la gramática o el tiempo de planificación presentes en la tarea son factores que pueden afectarla. Sin embargo, la fluidez rara vez se estudia en comparación con otras dimensiones y, cuando se hace, suele ser con relación a la comprensibilidad, ya que, por un lado, es la dimensión con la que mantiene una mayor relación y, por otro lado, ambas son dimensiones que tienen una fuerte presencia para el éxito de la comunicación (Derwing y otros 2007).

2.5. *La irritabilidad*

La irritabilidad se define como el grado de malestar y de rechazo que puede producir en el oyente la falta de dominio del hablante en cualquiera de las dimensiones estudiadas. Es por tanto una reacción emocional generada por las diferencias fónicas con respecto a la «norma» esperada, por las interrupciones, las interferencias, los malentendidos o los cortes en la comunicación; o bien por los esfuerzos «extras» que ha de realizar el oyente para comprender el mensaje, aunque también pueden intervenir factores no lingüísticos como el tema y el contexto de la conversación, la relación del oyente con el hablante, la voz y el aspecto físico de este último, o el estereotipo que se tiene de él o del grupo al que pertenece (Fayer y Krasinski 1987).

La falta de dominio de los contenidos lingüísticos que afecten a cualquiera de las anteriores dimensiones puede provocar reacciones en el interlocutor de muy diverso tipo, en muchos casos negativas. Algunas de estas reacciones pueden ir desde la impaciencia o la ansiedad, hasta la incomodidad o el desinterés, y afectan a la interacción que realiza el nativo, el cual puede decidir hablar más alto, más despacio, modificar su pronunciación, «desconectar» durante la conversación, alterar la comunicación cambiando de tema, o incluso interrumpirla evitando una futura interacción (Llurda 1995).

Algunos autores han considerado la inteligibilidad como la dimensión que genera niveles más altos de irritabilidad en comparación con otras dimensiones (Gallardo del Puerto y otros 2006), hasta el punto de que la irritabilidad se ha valorado como una propiedad inherente a la propia inteligibilidad en los casos en los que esta se produce (Ludwig 1982). Sin embargo, una buena inteligibilidad no garantiza necesariamente la ausencia de irritabilidad. Ciertos estudios han demostrado que niveles bajos de comprensibilidad (motivada por ejemplo por la conciencia de que el oyente tiene que invertir más tiempo en procesar la información)

(Munro y Derwing 1995b) y determinados grados de acento extranjero (Fayer y Krasinski 1987) pueden producir a su vez niveles muy elevados de irritabilidad, aunque el mensaje se entienda perfectamente.

3. MÉTODOS

3.1. *Participantes*

Las muestras orales se obtuvieron de ocho hablantes procedentes de Estados Unidos, todos universitarios, con edades comprendidas entre los 20 y los 22 años (media = 20,75), y de los cuales el 75% son mujeres.

Ninguno de ellos había vivido antes en un país hispanohablante y habían sido clasificados por los centros en los que estudiaban con un nivel avanzado de español (B2), aunque todos tenían distintos niveles de competencia fonética. En el momento de la obtención de las muestras todos llevan cuatro meses en España y viven con familias españolas. Ninguno de ellos ha recibido instrucción sobre pronunciación o corrección fonética.

Un total de 44 evaluadores nativos participó en el experimento: 15 expertos (profesores de español como L2 con más de cinco años de experiencia) y 29 no expertos (hablantes sin experiencia en la docencia ni formación universitaria en lingüística). El hecho de elegir a evaluadores nativos expertos y no expertos intenta reflejar el tipo de interlocutor en español que los hablantes de L2 pueden encontrar más frecuentemente en el contexto de inmersión en el que se encuentran. Por un lado, los profesores son los evaluadores habituales en el contexto del aula, mientras que los no expertos son los interlocutores en español más habituales fuera de ella. De este modo, los resultados ofrecidos por ambos grupos (y su comparación posterior) pueden arrojar datos interesantes sobre el modo en el que operan y se relacionan las cinco dimensiones de análisis dentro de cada grupo.

El grupo de expertos está formado por ocho hombres (53,3%) y siete mujeres (46,7%) que imparten clase de lengua española a extranjeros (fundamentalmente estadounidenses) en distintas instituciones académicas. Sus edades están equilibradas, y por tanto no hay diferencias significativas con $p > ,05$ (Chi2 corregido=0,000; 1 gl; N=15; $p > ,999$).

El 100% de los evaluadores expertos está familiarizado con el inglés y todos tienen también cierto grado de familiaridad con otras lenguas, fundamentalmente porque son las lenguas maternas de los estudiantes a los que enseñan. En cuanto a los años de experiencia como profesor de ELE, este grupo de expertos tiene una media (y también mediana) de 11 años ($11,07 \pm 5,38$) dentro de un rango de entre 5 y 20 años. Un 53,3% (8) posee formación en evaluación y solamente un 20% (3) la tiene en enseñanza de la pronunciación o de la corrección fonética.

Para la evaluación de las muestras por jueces no expertos se ha conseguido una muestra de 29 participantes: 16 mujeres (55,2%) y 13 hombres (44,8%). Para eliminar la edad como variable que pudiera influir en los resultados (Burda y otros 2003), se decidió considerar un rango de entre 30 y 50 años. Un 61,5% de los hombres (8 de 13) y un 56,2% de las mujeres (9 de 16) se encuentran en la franja de edad de los 30-40 años; el resto se sitúa en la banda de 40-50 años. Las diferencias no presentan significación con $p > ,05$ (Chi2 corregido=0,000; 1 gl; N=29; $p > ,999$). En total, algo más de la mitad de este grupo (el 58,6%) tiene entre 30-40 años.

Finalmente, un 58,6% (17 de 29) informa de que su nivel de inglés es alto (B2: C2) mientras que el otro 41,4% (12) refiere un nivel medio-bajo (A1: B1), lo que indica que todos los miembros del grupo tienen al menos un mínimo conocimiento de esta lengua.

3.2. *Procedimiento para la recogida de las muestras de los hablantes*

El experimento tuvo lugar en una habitación silenciosa en el edificio que Hamilton College tiene en Madrid. Las grabaciones se realizaron de forma individual. Todos los hablantes leyeron y firmaron un formulario de consentimiento y completaron un pequeño cuestionario con el investigador presente y disponible en un lugar discreto de la habitación para resolver sus dudas, pero sin interferir con su presencia.

El tipo de prueba usada para elicitación de las muestras orales fue la narración sobre un tema que resultara familiar a los estudiantes, en este caso sus últimas vacaciones. El investigador procuró crear un clima de confianza, de modo que la grabación incluyó un pequeño diálogo previo que no se tuvo en cuenta para la obtención posterior de la muestra. El hecho de utilizar esta prueba, y no otras, se debió a que fue considerada una técnica muy apropiada para conseguir un tipo de muestra oral que intentaba simular contextos de producción «natural», en los que un hablante expone a otro de forma detallada un tema con el que tiene mucha familiaridad.

Las ocho muestras se obtuvieron mediante una grabadora H1 Handy Recorder, a una frecuencia de muestreo de 16 bit / 44.1 kHz. Las muestras se editaron con el programa Audacity y se eligieron aproximadamente 60 segundos de la parte media de la muestra total del hablante. Todas las muestras se dispusieron de forma aleatoria en cada una de las tres sesiones de las que constaba el cuestionario. Con el fin de evitar saltos bruscos en los niveles de audio entre las muestras, se igualó la intensidad del audio de todas a 85 db (rango 84,6-86, 2).

3.3. *El cuestionario*

Debido a las dificultades encontradas para obtener un número significativo de jueces evaluadores que pudieran realizar el análisis perceptivo en

una misma sesión, se optó por diseñar dos cuestionarios en internet (uno para cada grupo de jueces) a través de la plataforma *google drive*. Todas las secciones venían precedidas de una explicación y de indicaciones para descargarse los tres archivos de audio (uno por cada sección) siguiendo un enlace. Con el fin de reducir los tiempos de descarga de los archivos (que eran muy pesados), se decidió comprimirlos a formato mp3 a 128kbps, y alojarlos en la plataforma <www.sugarsync.com> para facilitar su descarga.

Los cuestionarios estaban divididos en tres secciones con la siguiente estructura:

- *Sección 1:* transcripción de 24 muestras (tres por hablante) de aproximadamente ocho palabras (media de 8,25 palabras \pm 2,09 dentro de un rango de entre 5 y 12 palabras con mediana 8). Cada una de estas muestras procedía de la parte inicial, media y final del discurso total (aproximadamente 60 segundos) de cada hablante. Los evaluadores disponían de 20 segundos entre muestra y muestra, anunciadas mediante una voz que señalaba la numeración de la muestra y un sonido de campana. Las muestras venían precedidas de un ejercicio de práctica.
- *Sección 2:* evaluación en una primera audición de la comprensibilidad de las ocho muestras, precedidas de una primera muestra que servía de práctica, a través de una escala Likert de 9 niveles. Posteriormente, ya en una segunda audición, selección de las categorías que determinan los problemas de comprensibilidad. Todas estas categorías venían acompañadas de una breve explicación de sus características.
- *Sección 3:* evaluación en una primera audición del acento extranjero, la fluidez y la irritabilidad de las ocho muestras, precedidas de una primera muestra que servía de práctica, a través de escalas Likert de 9 niveles. Posteriormente, ya en una segunda audición, selección por parte de los jueces expertos de las categorías que determinan los problemas de acento extranjero. Todas estas categorías venían acompañadas de una breve explicación de sus características.

El orden de estas estructuras no es casual, y sigue una determinada lógica. La inteligibilidad se mide primero para garantizar el propio marco operativo y definitorio del constructo. No es un fenómeno escalar, sino binario (se entiende o no se entiende) y la objetividad de su método de medida se perdería ante «segundas oportunidades» que conllevarían la intervención de variables (marcos contextuales más amplios de tipo discursivo, efecto memoria) que ayudarían a reconstruir el mensaje. El hecho de incluir la comprensibilidad en segundo lugar tiene que ver con el propio diseño analítico que se ha realizado para esta dimensión. Siguiendo las investigaciones que señalábamos en la Sección 2.2, la comprensibilidad

(Sección 2 del experimento) está estructurada en cinco grandes bloques: pronunciación, fluidez, gramática, léxico y aspectos discursivos, mientras que el acento extranjero y la fluidez (Sección 3) no lo están.

Incluir el acento extranjero o la fluidez antes que la comprensibilidad habría supuesto el riesgo de que, al llegar a esta sección, los evaluadores tendieran a penalizar más los bloques correspondientes a pronunciación y fluidez analizados anteriormente, a pesar de enmarcarse en dimensiones distintas. Por otro lado, en este experimento la comprensibilidad tampoco se analiza junto con la inteligibilidad (es decir, no se hicieron a la vez, en una misma sección, cf. Derwing y Munro 1997), para impedir que los evaluadores asociaran sus notas Likert de comprensibilidad con su habilidad para transcribir las muestras analizadas, lo que provocaría que los evaluadores establecieran inconscientemente una relación directa entre ambas dimensiones, que es precisamente lo que se quiere evitar.

Debido a las características de los jueces y a la naturaleza de las dimensiones estudiadas, en este punto se decidió que los cuestionarios tuvieran niveles de introspección distintos para las dimensiones de comprensibilidad y de acento extranjero. Por ejemplo, en la Sección 2 (comprensibilidad), además de poner una nota, los evaluadores no expertos tienen que señalar en una segunda audición cuál(es) de los cinco grandes bloques que comentábamos (pronunciación, fluidez, gramática, léxico y aspectos discursivos) dificulta(n) la comprensibilidad. El grupo de expertos tiene que hacer lo mismo, pero en su caso señalando entre 28 subapartados distribuidos en esos cinco grandes bloques⁵. La razón de proceder de este modo estriba en que los expertos están acostumbrados a usar escalas analíticas de cierta amplitud en los procesos de evaluación. Por otro lado, existen estudios (Trofimovich e Isaacs 2012) que señalan importantes diferencias cuantitativas y cualitativas en los análisis realizados por jueces expertos y no expertos (siempre a favor de los primeros), por lo que se optó por diseñar un estudio analítico que pudiera aprovechar las ventajas y capacidades de este grupo frente a los no expertos. Esto mismo se realizó para el análisis del acento extranjero (Sección 3), donde el grupo de expertos debía señalar entre 12 variables⁶, frente a ninguna de los no expertos. Tales diferencias de metodología hacen que las comparaciones entre las notas asignadas por cada grupo a las categorías de comprensibilidad deban acogerse con la debida cautela, dado el distinto grado de profundidad analítica de las mediciones empleadas. Sin embargo, resultan

5. Los subapartados de esta categoría pueden verse en la Tabla 5. Cada uno de ellos viene acompañado de una breve explicación y se pide a los evaluadores que se familiaricen con ellos en el ejercicio de práctica antes de comenzar con las muestras.

6. Los subapartados de esta categoría pueden verse en la Tabla 6. Al igual que sucedía con el análisis de la comprensibilidad, todos los subapartados vienen acompañados de una breve explicación.

especialmente útiles para ver en qué se fija cada grupo desde un punto de vista perceptivo, así como qué aspectos concretos (dada la profundidad de los criterios analíticos manejados) señalan concretamente los expertos, los cuales no dejan de ser los evaluadores habituales y cualificados, en contextos de evaluación académicos.

La fluidez y la irritabilidad se miden también a través de escalas Likert de 9 niveles, y el procedimiento se realiza de idéntica forma en los dos grupos. La elección de una escala Likert de nueve niveles y no de otro número se ha hecho siguiendo dos consideraciones: por un lado las escalas de nueve niveles son las que han alcanzado mayor implantación en los últimos años; por otro lado, aunque no existe la escala perfecta, las escalas demasiado restrictivas pueden ser problemáticas para mostrar distinciones, por lo que, en la línea de lo que señalan Munro y Derwing a propósito de la habilidad del oyente para evaluar el acento extranjero, consideramos también que resulta mejor «to overestimate the listener's ability [...] than to underestimate it» (1994, p. 259).

Por último, y en relación con este punto, decidimos no utilizar hablantes nativos en las muestras (solo hablantes de español como L2) para reducir el riesgo de que los evaluadores reservaran precisamente los extremos a estos hablantes y la escala perdiera con ello nivel de distinción. También porque, en los contextos de evaluación, este es el tipo de situación más frecuente que los evaluadores (al menos los expertos) encuentran y porque los hablantes nativos, en general, tienen ya interiorizada una percepción de lo que constituye en sí mismo el acento extranjero (Trofimovich e Isaacs 2012).

4. RESULTADOS

El análisis estadístico se ha realizado mediante la aplicación SPSS Statistics 20, fijando un nivel de significación del 5% ($p < 0,050$). Con el fin de facilitar el seguimiento, se muestran a continuación los resultados estructurados por dimensión, comparando los grupos de evaluadores cuando así procede.

4.1. *La inteligibilidad*

El análisis comienza verificando si se detectan diferencias que sean estadísticamente significativas entre las diferentes muestras (tres por hablante) y entre cada uno de los evaluadores de cada grupo, en lo que se refiere al porcentaje de palabras que entienden correctamente⁷. Para ello se empleó un Análisis de Varianza (ANOVA) de doble clasificación con

7. Ningún juez señaló en el apartado indicado problemas de tiempo a la hora de transcribir las muestras, aspecto que hubiera supuesto su exclusión de los análisis realizados.

medidas repetidas, con estimación del tamaño del efecto mediante Eta^2 , tanto para el grupo de expertos como para el de no expertos.

Hablaante	No experto		Experto		ORDEN que ocupa el hablaante, según:	
	%medio	D.E.	%medio	D.E.	NO EXPERTOS	EXPERTOS
H-1	0,97	0,04	0,94	0,04	2º	4º
H-2	0,97	0,06	0,98	0,05	2º	1º
H-3	0,72	0,17	0,64	0,13	8º	8º
H-4	0,86	0,11	0,78	0,11	7º	7º
H-5	0,98	0,03	0,93	0,07	1º	5º
H-6	0,97	0,07	0,98	0,04	2º	1º
H-7	0,96	0,08	0,97	0,04	5º	3º
H-8	0,94	0,08	0,91	0,08	6º	6º

TABLA 1. Medias totales y desviaciones estándar del porcentaje de aciertos por hablaante; comparativa de rangos entre jueces expertos/no expertos.

Debido a que se apreciaron diferencias significativas entre las tres muestras evaluadas en cada hablaante, así como entre los propios evaluadores, se decidió contrastar la posible existencia de diferencias en función del grupo de evaluadores (expertos/no expertos) en cada hablaante por separado y en cada muestra, por tanto en el total de las 24 muestras utilizadas. Los resultados mostraron que en la mayoría de las frases (en 22 de las 24, un 91,7%), no existen diferencias que puedan ser justificadas estadísticamente con $p > ,05$, y en las dos donde sí las había (con $p < ,05$), los tamaños del efecto eran leves (.116 y .158), lo que permite concluir que, en general, no ha habido diferencia en los resultados de estas categorías entre un grupo de jueces y otro.

Por ello, a pesar de las divergencias entre unos jueces y otros observadas internamente en su propio grupo y visto que la mayoría de los valores estaban por encima del 90% de aciertos con poca variabilidad, se decidió resumir las diferentes medias en un solo valor medio total por hablaante. La Tabla 1 presenta estas medias totales comparadas entre el grupo de expertos y no expertos, así como el rango que le correspondería a cada hablaante en una clasificación en cada grupo. La semejanza de ambas clasificaciones apoya claramente la ausencia de diferencias significativas en la evaluación de la inteligibilidad entre jueces expertos y no expertos. Como medida del grado de coincidencia entre ambas clasificaciones realizadas con los ocho hablaantes, se empleó el coeficiente de correlación de Spearman obteniendo un valor de $r = 0,687$ (elevada coincidencia), que es significativo con $p < ,05$ ($p = ,030$), lo que demuestra el muy buen grado de coincidencia entre las clasificaciones resultantes de los juicios de los dos grupos.

Por otro lado, se procedió a comparar la anterior variable de porcentajes de aciertos con otros factores que podrían ser causa de diferencias, como el sexo (Schairer 1992) o la edad (Burda y otros 2003). En ninguna de las dos variables se detectaron diferencias que alcanzaran significación estadística con $p > ,05$ y unos tamaños del efecto mediante η^2 muy leves (entre ,000 y ,068) en ninguna de las 24 muestras habladas, por lo que no podemos decir que los porcentajes de acierto obtenidos dependan del sexo o la edad de los jueces⁸. Asimismo, se estudió también la relación entre el número de palabras de cada una de las muestras con el porcentaje medio de aciertos obtenido para cada una de ellas, concluyendo que la longitud en palabras de cada muestra usada en este estudio no afectó tampoco al porcentaje de aciertos de inteligibilidad.

Por último, un análisis particularmente interesante consiste en conocer la relación de los porcentajes reales⁹ de acierto y la percepción subjetiva que de ello tuvo cada uno de los jueces, con el fin de determinar si hay correspondencia entre ellos¹⁰, tal y como se observa en la Tabla 2. Este autor no tiene conocimiento de estudios que hayan comparado estos dos tipos de medidas (una impresionista y la otra metodológicamente científica) hasta la fecha. Los resultados pueden aportar pistas tanto sobre la validez de algunos estudios que han medido la inteligibilidad en forma de escalas por niveles (Fayer y Krasinski 1987) o continuas (Isaacs 2008), ambas de tipo impresionista, como sobre el «desfase» real entre lo que los hablantes dicen que entienden, y lo que verdaderamente entienden.

Los números marcados en negrita en la Tabla 2 precisan los jueces cuya percepción subjetiva coincide con el porcentaje real de acierto en la prueba. Ha resultado en 14, un 31,8%. Los que se encuentran a la izquierda de esa diagonal de la tabla (9 de 44, el 20,45%) son los que subjetivamente creen haber tenido menos inteligibilidad que la que tuvieron realmente, es decir, que se infravaloran. Y los que se encuentran por encima (21, el 47,7%) son los que se sobrevaloran creyendo que su inteligibilidad es mayor. Por tanto, el grado de concordancia obtenido es muy bajo (Kappa = 0,088) y no alcanza significación estadística con $p > ,05$ ($p = ,274$). Estos resultados pueden verse resumidos aquí en forma de gráfico:

8. Se optó por trabajar con todos los jueces juntos en única muestra (N=44) dada la ausencia de diferencias entre ellos que se acaba de probar. Se emplea la misma estrategia estadística: Anova de 1 factor.

9. Se procedió a calcular una media total del porcentaje de aciertos para cada uno de los jueces, promediando por tanto las 24 frases en un único valor. A partir de ese valor, se categoriza esta variable del porcentaje medio total/real de aciertos, de la misma forma en que se les pidió una evaluación subjetiva de lo que habían entendido en rango de porcentaje.

10. Se utilizó χ^2 de tablas de contingencia como estadística de significación y la Kappa de Cohen como método de estimación de la intensidad de la concordancia entre una y otra variable. Dada la escasa diferencia entre los dos grupos de jueces, el estudio se realizó con la muestra total de 44 jueces.

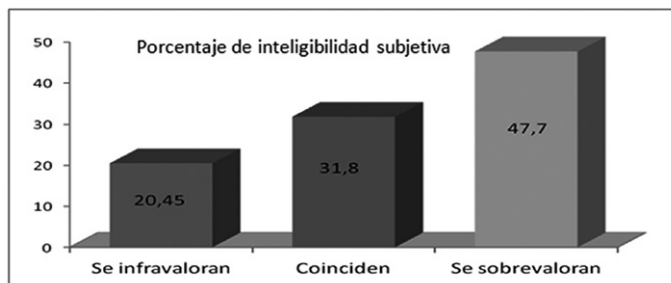


GRÁFICO 1. Porcentaje de inteligibilidad subjetiva.

		INTELIGIBILIDAD SUBJETIVA						Total	
		60%-64%	75%-79%	80%-84%	85%-89%	90%-94%	95%-99%		
Categorías del porcentaje medio total de aciertos	75%-79%	Nº jueces	0	0	0	1	0	0	1
		% del total	0,0%	0,0%	0,0%	2,3%	0,0%	0,0%	2,3%
	80%-84%	Nº jueces	1	0	1	0	1	0	3
		% del total	2,3%	0,0%	2,3%	0,0%	2,3%	0,0%	6,8%
	85%-89%	Nº jueces	0	1	1	3	4	6	15
		% del total	0,0%	2,3%	2,3%	6,8%	9,1%	13,6%	34,1%
	90%-94%	Nº jueces	0	0	1	1	5	9	16
		% del total	0,0%	0,0%	2,3%	2,3%	11,4%	20,5%	36,4%
95%-99%	Nº jueces	0	0	0	1	3	5	9	
	% del total	0,0%	0,0%	0,0%	2,3%	6,8%	11,4%	20,5%	
Total		Nº jueces	1	1	3	6	13	20	44
		% del total	2,3%	2,3%	6,8%	13,6%	29,5%	45,5%	100%

TABLA 2. Cruce de la inteligibilidad subjetiva con el porcentaje medio real de aciertos.

4.2. La comprensibilidad

Como se comentó en la Sección 3.3, la evaluación de la comprensibilidad se hizo a través de una escala Likert de 9 niveles: 1: extremadamente fácil de entender; 9: extremadamente difícil de entender. La Tabla 3 presenta las medias totales y las desviaciones estándar de la comprensibilidad de cada hablante en cada grupo de jueces, así como el orden, establecido según las valoraciones medias (por ser esta herramienta estadística más sensible que la moda), en que cada uno de estos grupos sitúa a los hablantes. Como puede apreciarse, en los dos grupos el orden asignado a los hablantes es muy similar al obtenido en inteligibilidad, por lo que podemos intuir que entre ambas dimensiones se establece una estrecha relación.

Existe además una clara coincidencia entre las clasificaciones de expertos y no expertos con respecto a aquellos hablantes con una peor comprensibilidad (los cuatro últimos). Se estimó aquí la correlación entre ambas clasificaciones mediante Spearman, obteniendo una elevada coincidencia ($r=,787$) altamente significativa con $p<,01$ ($p=,009$). Del mismo modo, la fiabilidad alcanzada por las valoraciones de cada grupo según «alfa» de Cronbach es elevada ($,850$ para los expertos y $,869$ para los no expertos) y altamente significativa con $p<,001$ ($p=,000$).

Habla nte	Experto		NO Experto		ORDEN que ocupa el habla nte, según:	
	Media	D.E.	Media	D.E.	EXPERTOS	NO EXPERTOS
H-1	3,27	1,71	2,17	1,51	2º	3º
H-2	3,33	1,99	1,90	1,32	3º	1º
H-3	5,40	1,35	4,14	1,90	8º	8º
H-4	4,73	1,67	2,83	1,39	7º	7º
H-5	2,33	0,90	2,31	1,49	1º	4º
H-6	3,47	1,30	1,93	1,25	4º	2º
H-7	3,73	1,44	2,38	1,08	5º	5º
H-8	4,47	1,64	2,76	1,57	6º	6º

TABLA 3. Medias totales y desviaciones estándar de la comprensibilidad de cada habla nte; comparativa de rangos entre jueces expertos/no expertos.

Como puede verse en la Tabla 3, en todos estos habla ntes la media de los valores Likert otorgados por los no expertos es menor que la media de los valores de los expertos, por lo que a la vista de los resultados se puede afirmar que para los jueces no expertos los habla ntes presentan mejor grado de comprensibilidad que para los jueces expertos. Para el contraste comparativo entre sí de los valores dados por los dos grupos se volvió a emplear (tal y como se había hecho en el apartado de inteligibilidad) el test estadístico Anova de 1 factor. Los resultados mostraron cómo en todos los habla ntes, excepto en uno (H-5), las diferencias alcanzaban significación estadística al menos con $p<,05$, e incluso en algunos casos hasta con $p<,01$. En los casos en los que la diferencia tiene menor significación el tamaño del efecto mediante valor de Eta^2 es leve, en torno a $,100$; pero en los casos donde hay alta significación el tamaño del efecto ya es moderado, en el entorno de $,250$. En la valoración global de la muestra de habla ntes como conjunto la diferencia es altamente significativa con $p<,001$, y el tamaño del efecto moderado ($,263$).

Respecto a las distintas categorías marcadas en el apartado de la comprensibilidad por expertos y no expertos en una segunda audición, conviene recordar que ambos grupos debían señalar una o varias categorías entre pronunciación, gramática, vocabulario, fluidez y aspectos discursivos

(cinco en total), o dejar la respuesta en blanco si no encontraban ninguna categoría que impidiera la comprensibilidad. El grupo de expertos tenía además que profundizar eligiendo entre alguna de las 28 subcategorías que aparecían distribuidas entre las cinco categorías generales.

Para las comparaciones entre grupos de jueces expertos y no expertos se han realizado porcentajes medios de cada una de las cinco categorías para cada uno de los ocho hablantes, calculando posteriormente el valor medio total de todos ellos por categoría (ver Tabla 4). En el caso de los jueces expertos, los porcentajes de cada categoría se han obtenido a partir de la suma de los porcentajes medios de las subcategorías seleccionadas para cada hablante, calculando después, de forma análoga, el valor medio total de todos ellos.

Como puede observarse en la Tabla 4, los problemas léxicos y gramaticales tienen mayor incidencia sobre la comprensibilidad para los jueces expertos que los no expertos. Sin embargo, para la pronunciación, los juicios de los no expertos alcanzan un promedio de 53,5%, que es superior no

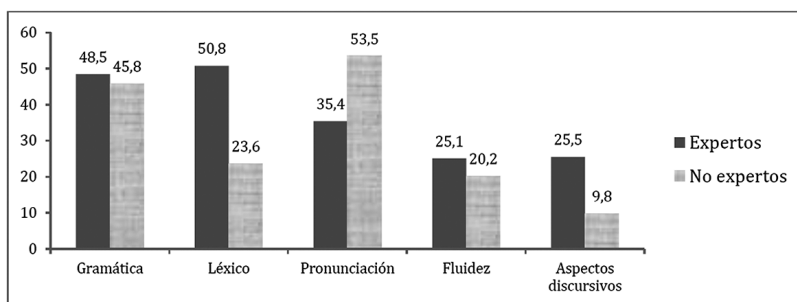


TABLA 4. Factores que inciden en la comprensibilidad para jueces expertos y no expertos.

solo al 35,5% de media de los expertos, sino incluso a todos los promedios de las nueve variables analizadas por ellos (véase Tabla 5). Por lo tanto, puede decirse que los no expertos son más sensibles a los problemas de pronunciación como factor que dificulta la comprensibilidad y que son los aspectos gramaticales y léxicos los que más influyen entre los expertos.

Estos resultados pueden explicarse por el hecho de que los evaluadores expertos tienen un mayor grado de familiaridad con este tipo de hablantes que tienen el español como L2 y la pronunciación como rasgo prominente no es, por tanto, un factor que los afecte tanto como al grupo de no expertos. A su vez, el hecho de que los jueces expertos encuentren los mayores problemas en los aspectos gramaticales y léxicos puede derivarse de las propias prácticas de enseñanza que enfatizan estas competencias, lo que les llevaría a ser más «intolerantes» con este tipo de errores que con otros que no tienen tanto peso en el aula (Barrera Pardo 2004). En lo

que se refiere a la fluidez, esta no es una elección recurrente por parte de los no expertos y, aunque el porcentaje comparativo es mayor, desciende también notablemente el porcentaje de jueces expertos que señalan este problema, algo que también sucede en los aspectos discursivos.

Por su parte, la Tabla 5 muestra de una forma detallada los valores de las subcategorías señalados dentro de cada una de estas cinco categorías por parte de los jueces expertos. Al igual que en la Tabla 4, estos valores indican el porcentaje de jueces que señalan estos aspectos en la media resultante de los ocho hablantes. Una cuestión que llama la atención es que en todas las categorías hay uno o dos elementos que destacan por encima del resto, lo que parece sugerir que en los procesos de evaluación de la comprensibilidad pueden existir aspectos lingüísticos más sensibles que otros, a los que los evaluadores dan más importancia o prestan más atención. Valorar en qué medida estos resultados responden a patrones subjetivos e internos de los evaluadores, o a juicios que se corresponden fielmente con la producción real de los hablantes, solo puede hacerse mediante un análisis estadístico que considere el material lingüístico y acústico de las muestras, aspecto que deja abiertas las puertas a nuevos y fértiles campos de estudio. En esta investigación, los aspectos más señalados por los expertos son la concordancia y la morfología verbal en la categoría de gramática (con cinco y tres hablantes respectivamente que tienen problemas en estos elementos según al menos el 80% de los jueces); la confusión léxica en la categoría de léxico (con tres hablantes donde al menos el 80% de los jueces hace esa valoración); las consonantes y vocales en la categoría de pronunciación; los falsos comienzos, reparaciones, reformulaciones y vacilaciones para la categoría de fluidez; y la profundidad discursiva y la cohesión como los dos elementos más problemáticos en la categoría de aspectos discursivos.

Con el fin de buscar diferencias estadísticamente significativas entre las valoraciones de los dos grupos, se ha empleado en cada una de las cinco categorías el test χ^2 de independencia y se han evaluado incoherencias y variabilidades, que en casi todos los casos (pues no es así en la categoría de fluidez o en algunas subcategorías, como las de consonantes y vocales) resultan estadísticamente significativas¹¹.

11. Comparando los porcentajes de problemas considerados para cada una de las cinco categorías por unos jueces y otros, así como los elementos para cada categoría señalados por los expertos, se han encontrado diferencias que son: a) estadísticamente significativas con $p < ,01$ ($\chi^2 = 21,87$; 7 gl; $p = ,003$) para la gramática; b) altamente significativas con $p < ,001$ ($\chi^2 = 46,68$; 7 gl; $p = ,000$) para el léxico; c) no significativas con $p > ,05$ ($\chi^2 = 12,09$; 7 gl; $p = ,098$) al comparar los porcentajes de los problemas señalados por los no expertos para la pronunciación con los valores de las subcategorías de consonantes y vocales señalados por los expertos; y sin embargo altamente significativas con $p < ,001$ ($\chi^2 = 21,79$; 7 gl; $p = ,003$) al comparar los porcentajes del resto de elementos con la valoración global dada para esa categoría por los no expertos; d) no estadísticamente significativas con $p > ,05$ ($\chi^2 = 11,24$; 7 gl; $p = ,128$) para la fluidez, aunque sí se

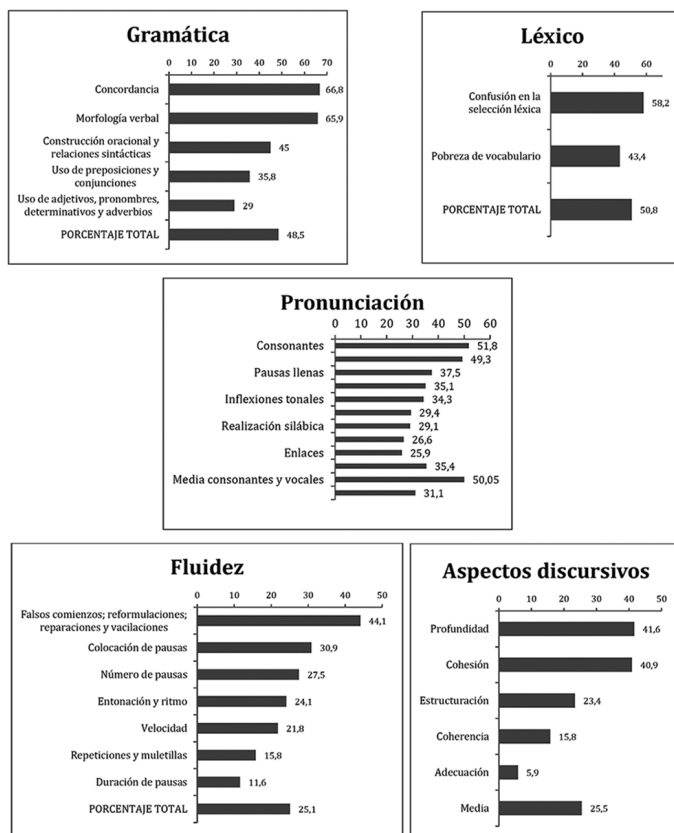


TABLA 5. Valores de las subcategorías que inciden en la comprensibilidad señalados por los jueces expertos.

4.3. El acento extranjero

Al igual que en la sección de comprensibilidad, la evaluación de esta dimensión se hizo a través de una escala Likert de 9 niveles: 1: sin acento extranjero; 9: con acento extranjero extremadamente marcado. La Tabla 6 presenta las medias totales y las desviaciones estándar del acento extranjero de cada hablante en cada grupo de jueces, así como el orden, establecido según las valoraciones medias en que cada uno de estos grupos sitúa a los

compara solamente el elemento más señalado, el de falsos comienzos, reformulaciones, reparaciones y vacilaciones con los juicios de los no expertos, sí que aparecen diferencias que son claramente significativas con $p < ,01$ ($\chi^2 = 18,75; 7 \text{ gl}; p = ,009$) que indican que los jueces expertos detectan más este problema de fluidez en concreto que los juicios holísticos de los no expertos; y significativas con $p < ,01$ ($\chi^2 = 18,75; 7 \text{ gl}; p = ,009$) para los aspectos discursivos.

hablantes. Como puede apreciarse, existe una destacada coincidencia entre las clasificaciones de expertos y no expertos, especialmente con respecto a los hablantes que se sitúan en los extremos (concretamente los tres primeros y los dos últimos). También en esta tabla se puede observar cómo en general las valoraciones dadas por los jueces no expertos tienden a ser más bajas que las de los expertos, lo cual muestra que puntúan más hacia el extremo señalado como ausencia de acento extranjero. Con el fin de realizar un contraste comparativo de los dos grupos según su valoración en la escala de Likert, se empleó nuevamente un test Anova de 1 factor, que mostró que no se alcanza significación estadística con $p > ,05$ y un tamaño del efecto leve ($,037$) en la escala η^2 . Por tanto podemos afirmar que los juicios emitidos sobre el grado de acento extranjero son similares en ambos grupos de jueces.

Por otro lado, la fiabilidad alcanzada por las valoraciones de cada grupo según «alfa» de Cronbach es elevada ($,821$ para los expertos y $,790$ para los no expertos), y altamente significativa con $p < ,001$ ($p = ,000$). También se estimó aquí la correlación entre ambas clasificaciones mediante Spearman, obteniendo una elevadísima coincidencia ($r = ,976$) muy significativa con $p < ,001$ ($p = ,000$).

Finalmente, la Tabla 7 muestra en detalle los aspectos valorados dentro de la categoría de acento extranjero por los jueces expertos, indicando el porcentaje de jueces que ha considerado cada uno de esos aspectos como influyentes en el acento extranjero. Resulta destacable la inclinación que los evaluadores expertos muestran, una vez más, por los aspectos segmentales, tal y como ocurría para la categoría «pronunciación» en el análisis sobre la comprensibilidad de las muestras.

Habla nte	Experto		NO Experto		ORDEN que ocupa el habla nte, según:	
	Media	D.E.	Media	D.E.	EXPERTOS	NO EXPERTOS
H-1	5,00	1,81	4,34	1,80	2º	2º
H-2	3,60	2,13	3,48	2,03	1º	1º
H-3	8,00	1,07	7,76	1,30	8º	8º
H-4	6,13	1,51	6,10	1,78	5º	6º
H-5	7,13	1,19	6,97	1,57	7º	7º
H-6	5,87	1,85	4,86	1,64	4º	4º
H-7	5,07	1,75	4,62	1,97	3º	3º
H-8	6,33	1,40	5,41	1,92	6º	5º

TABLA 6. Medias totales y desviaciones estándar grado de acento extranjero de cada habla nte; comparativa de rangos entre jueces expertos/no expertos.

Elemento	Habla nte (los valores indican % de jueces)								Total		
	H-1	H2	H-3	H-4	H-5	H-6	H-7	H-8	Media	D.E.	Rango
Vocales	67	53	73	73	67	67	60	47	63,4	9,3	1º
Consonantes	53	80	73	73	73	47	53	27	60,0	18,0	2º
Pausas llenas	40	20	53	73	40	53	33	53	45,6	16,0	3º
Rango tonal	40	13	47	20	60	27	60	47	39,3	17,7	4º
Realización acentual	53	33	47	20	40	20	33	53	37,4	13,3	5º
Inflexiones tonales	13	27	40	80	33	40	20	40	36,6	20,2	6º
Posición acentual	40	7	27	27	33	20	47	87	36,0	23,9	7º
Realización silábica	53	0	53	20	60	13	40	27	33,3	21,6	8º
Colocación pausas	13	20	33	27	40	40	33	40	30,8	10,1	9º
Agrupaciones	7	27	27	27	27	33	33	27	26,0	8,1	10º
Enlaces	7	7	27	13	33	27	20	13	18,4	9,9	11º
Velocidad	7	7	27	13	13	27	33	13	17,5	10,0	12º

TABLA 7. Acento extranjero. Aspectos que influyen, según los expertos.

En este apartado, vocales y consonantes vuelven a ser los más destacados con promedios a partir del 60%, y superando en varios hablantes el 70%. Todos los demás aspectos se quedan por debajo del 50% en promedio.

En la línea de lo que señalaba Odisho 2003, p. 53, este tipo de resultados puede responder a concepciones «simplistas» imperantes en el aula, que conciben el habla como una categorización de sonidos divididos fundamentalmente en dos grupos: vocálicos y consonánticos. Sería necesario definir entonces si esta percepción se corresponde con los resultados acústicos de las muestras o si, por el contrario, es más propia de una ausencia de formación en fonética que derive los análisis hacia aquellos aspectos más conocidos por los evaluadores.

4.4. La fluidez

Al igual que en las secciones de comprensibilidad y acento extranjero, la evaluación de esta dimensión se hizo a través de una escala Likert de 9 niveles: 1: gran fluidez; 9: enormes dificultades de fluidez. Como puede observarse en la Tabla 8, los dos grupos coinciden en indicar a los mismos tres primeros hablantes como los que mayor grado de fluidez presentan en español, además de que los jueces no expertos tienden a dar valoraciones de menor valor, esto es, a indicar una mayor fluidez en general de la que dieron los expertos. El grado de fiabilidad de ambos grupos es elevado (alfa ,831 para los expertos y ,811 para los no expertos, en ambos casos muy

significativos con $p < ,001$ ($p = ,000$). Al correlacionar comparativamente ambas clasificaciones mediante Spearman se ha obtenido un coeficiente elevado ($r = ,575$), aunque algo menor que los anteriores, y que está en el límite de la significación con una p ligeramente superior a $,05$ ($p = ,068$). Por otra parte, en el análisis comparativo de ambos grupos se ha observado que en los seis primeros hablantes las diferencias no han alcanzado significación estadística con $p > ,05$. A pesar de ello, en la media total acumulativa de todos ellos hay una diferencia que tiene significación con $p < ,05$ y un tamaño del efecto leve mediante valor de η^2 ($,128$). Por tanto estamos ante una situación de cierta indefinición: por un lado, aunque los jueces no expertos puntúan en niveles más bajos (dando por tanto notas de más fluidez) no llega a haber en la mayoría de los casos diferencias significativas, pero puesto que en algunos hablantes las diferencias son más grandes (H-7 y H-8), a nivel grupal sí se alcanza significación.

Hablaante	Experto		NO Experto		ORDEN que ocupa el hablante, según:	
	Media	D.E.	Media	D.E.	EXPERTOS	NO EXPERTOS
H-1	3,47	1,55	2,72	1,00	2º	2º
H-2	2,67	1,50	2,24	1,70	1º	1º
H-3	5,00	1,51	4,03	1,88	6º	6º
H-4	4,87	1,64	4,24	1,94	5º	7º
H-5	4,07	1,58	3,31	1,67	3º	3º
H-6	4,60	1,72	4,45	1,68	4º	8º
H-7	5,00	1,77	3,45	1,62	6º	4º
H-8	5,13	1,73	3,52	1,40	8º	5º

TABLA 8. Medias totales y desviaciones estándar del grado de fluidez de cada hablante; comparativa de rangos entre jueces expertos/no expertos.

4.5. La irritabilidad

La dimensión de la irritabilidad se evaluó también a través de una escala Likert de 9 niveles: 1: ninguna irritabilidad; 9: enorme irritabilidad. Como puede apreciarse en la Tabla 9, existe un muy elevado grado de coincidencia entre las clasificaciones ordinales en las que cada grupo sitúa a los hablantes, especialmente en los extremos (los dos primeros y los dos últimos). Por otra parte, los jueces no expertos tienden a puntuar en niveles más bajos, indicando menor grado de irritabilidad que los jueces expertos. El grado de fiabilidad de las valoraciones de los dos grupos es elevado (alfa: $,921$ para los expertos, la mayor de todas las obtenidas, y $,817$ para los no expertos, en ambos casos muy significativos con $p < ,001$ ($p = ,000$)).

Como era de esperar, al correlacionar entre sí ambas clasificaciones completas con el coeficiente de Spearman, se ha obtenido un valor muy elevado ($r = ,850$) y significativo con $p < ,01$ ($p = ,004$) que indica la gran coincidencia entre ambos órdenes. En la comparación de las valoraciones de ambos grupos se ha observado que existen diferencias altamente significativas con $p < ,001$ y tamaño del efecto mediante valor de η^2 moderado ($,222$) en la media total acumulativa de todos los hablantes. Aunque en tres de ellos (H-2, H-3 y H-4) no ha habido diferencias estadísticamente significativas con $p > ,05$, en el resto, que son mayoría, sí que las ha habido, y en casi todos ellos han sido significativas con $p < ,01$ e incluso con $p < ,001$ y con tamaños del efecto mediante η^2 moderados, en el entorno de $,200$.

Habla nte	Experto		NO Experto		ORDEN que ocupa el habla nte, según:	
	Media	D.E.	Media	D.E.	EXPERTOS	NO EXPERTOS
H-1	3,73	2,02	2,00	1,25	2º	2º
H-2	2,53	1,73	1,69	1,23	1º	1º
H-3	5,13	2,30	4,45	2,44	8º	8º
H-4	4,73	2,31	3,48	2,23	7º	7º
H-5	4,07	2,25	2,66	2,99	3º	6º
H-6	4,13	2,32	2,07	0,92	4º	3º
H-7	4,13	2,03	2,34	1,68	4º	4º
H-8	4,60	2,06	2,38	1,59	6º	5º

TABLA 9. Medias totales y desviaciones estándar del grado de irritación de cada habla nte; comparativa de rangos entre jueces expertos/no expertos.

4.6. *Interrelaciones entre las cinco dimensiones*

El procedimiento que se siguió para homogeneizar la forma de presentar los datos de las variables fue recoger las clasificaciones ordinales realizadas en cada una de esas supracategorías. La Tabla 10 presenta la matriz de intercorrelaciones¹² obtenidas para el grupo de jueces expertos. Se ha encontrado que el acento extranjero no correlaciona significativamente con ninguna de las demás supracategorías ($p > ,05$), excepto con irritabilidad. El resto de coeficientes son todos significativos, al menos con $p < ,05$. Inteligibilidad y comprensibilidad tienen la relación de mayor intensidad ($r = ,976$), altamente significativa ($p = ,000$), lo que indica que las dos clasificaciones ordinales son muy semejantes en todos sus rangos. Esto

12. Con estas variables se procedió a calcular las correlaciones mediante el coeficiente de Spearman, más adecuado al tipo y número de datos que el coeficiente de Pearson.

parece pues avalar los estudios que han señalado la relación más fuerte entre ambas dimensiones (Munro y Derwing 1999).

		Comprensibilidad	Acento Extranjero	Fluidez	Irritabilidad
Inteligibilidad	Coefficiente de correlación	,976	,390	,724	,859
	Valor p-sig.	,000**	,170 ^{NS}	,021*	,003**
Comprensibilidad	Coefficiente de correlación		,429	,755	,898
	Valor p-sig.		,145 ^{NS}	,015*	,001**
Acento Extranjero	Coefficiente de correlación			,587	,731
	Valor p-sig.			,063 ^{NS}	,020*
Fluidez	Coefficiente de correlación				,825
	Valor p-sig.				,006**

NS = no significativo ($p > ,050$) * Significativo al 5% ** Altamente Significativo al 1%.

TABLA 10. Correlaciones entre todas las supracategoría, según las clasificaciones de los expertos.

		Comprensibilidad	Acento Extranjero	Fluidez	Irritabilidad
Inteligibilidad	Coefficiente de correlación	,922	,802	,323	,886
	Valor p-sig.	,001**	,008**	,217 ^{NS}	,002**
Comprensibilidad	Coefficiente de correlación		,762	,476	,905
	Valor p-sig.		,014*	,116 ^{NS}	,001**
Acento Extranjero	Coefficiente de correlación			,548	,952
	Valor p-sig.			,080 ^{NS}	,000**
Fluidez	Coefficiente de correlación				,548
	Valor p-sig.				,080 ^{NS}

NS = no significativo ($p > ,050$) * Significativo al 5% ** Altamente Significativo al 1%.

TABLA 11. Correlaciones entre todas las supracategorías, según las clasificaciones de los no expertos.

La comprensibilidad también tiene una relación clara, y altamente significativa ($p = ,001$), con la irritabilidad ($r = ,898$), según la cual los sujetos con menor irritabilidad serían los primeros en comprensibilidad. Algo muy similar ocurre con inteligibilidad e irritabilidad ($r = ,859$, con $p = ,003$), siendo este triángulo de variables las que más intensamente correlacionan entre sí. Esto se corresponde también con los estudios que señalan a los errores que impiden la comunicación, y a los que la dificultan, como los

principales causantes de la irritación del oyente (Fayer y Krasinski 1987; Gallardo del Puerto y otros 2006). La correlación de fluidez con irritabilidad también es elevada ($r=,825$; $p=,006$), pero ya no lo es tanto con las otras dos, comprensibilidad ($r=,755$; $p=,015$) e inteligibilidad ($r=,724$; $p=,021$).

La Tabla 11 es la matriz de correlaciones obtenidas con los juicios de los no expertos, mostrando unos resultados sustancialmente diferentes. En este caso es la fluidez la variable que no correlaciona significativamente con ninguna de las demás ($p>,05$). Por el contrario acento extranjero sí correlaciona con las demás, especialmente con irritabilidad ($r=,952$; $p=,000$), siendo esta la que se ha encontrado con mayor intensidad de todas las que se han calculado. Estas serían por tanto las principales diferencias con respecto a los expertos.

Inteligibilidad y comprensibilidad mantienen una elevada relación ($r=,922$; $p=,001$). Asimismo se mantienen, aunque con una ligera mayor intensidad, las relaciones entre inteligibilidad e irritabilidad ($r=,886$; $p=,002$), y entre comprensibilidad e irritabilidad ($r=,905$; $p=,001$).

5. DISCUSIÓN Y CONCLUSIONES

Los resultados obtenidos permiten probar de modo directo que estas dimensiones mantienen entre sí distintas relaciones que pueden variar en función del grupo evaluador. Con todo, la inteligibilidad y la comprensibilidad aparecen en todos los casos como las dimensiones con mayor correlación y siempre están asociadas con la irritabilidad. El acento extranjero, y en mucha mayor medida la fluidez, ofrecen relaciones menos sólidas.

El acento extranjero es el constructo que marca más la diferencia en las valoraciones perceptivas entre jueces expertos y no expertos. Para los no expertos, es la dimensión que más se correlaciona con la irritabilidad, y la categoría que más incide en la comprensibilidad. Aunque la inmensa mayoría de los evaluadores de este grupo afirma conocer la L1 de los hablantes, en este punto no hay mucha diferencia con los expertos, y, sin embargo, estos últimos no señalan la pronunciación como factor que dificulte la comprensibilidad, ni tampoco ofrecen correlaciones tan altas con la irritabilidad. Todo ello parece sugerir que el grado de exposición al acento extranjero (que es el que tienen en este estudio los evaluadores expertos) ha podido ser un factor mucho más determinante para mejorar los niveles de comprensibilidad. Se necesitarían, pues, más estudios que analizaran las variables no lingüísticas que intervienen en las relaciones que se establecen entre estas dimensiones.

Este punto nos lleva a señalar también otras vías de investigación. Una de ellas consistiría en realizar análisis que permitan poner en relación, con respecto al español, las notas de los evaluadores con las características acústicas de las muestras de habla. Esto nos permitiría identificar científicamente los aspectos lingüísticos que más inciden en cada una de estas

dimensiones, lo que sería muy útil para elaborar tanto planes de actuación didácticos como escalas y descriptores de evaluación dirigidos a medir aquellos aspectos que, en primer lugar, impiden y dificultan la comunicación. Por otra parte, los análisis de este tipo nos podrían proporcionar también interesantes resultados sobre el grado de correspondencia entre los juicios perceptivos de los evaluadores y los datos «reales», de naturaleza acústica y descriptiva, que ofrecen estas muestras. Serían especialmente útiles en contextos de evaluación y nos permitirían entender mejor qué es lo que determina una reacción concreta. No hay que olvidar tampoco que, en muchas ocasiones, los prejuicios y estereotipos que los hablantes tienen acerca de una particular L2 pueden activar expectativas sobre determinadas características de la muestra que quizá ni siquiera están presentes en la señal sonora (Kang y Rubin 2009). La meta última, por tanto, sería informar a los profesores y a los hablantes sobre los objetivos lingüísticos a los que deberían enfocar su atención, y conocer con mayor precisión las variables que influyen y determinan la percepción y la relación de estos constructos.

Como hemos visto, el acento se ha convertido, pues, en un fértil campo de estudio en el que se han abierto (y se están abriendo) líneas de investigación muy variadas. Algunas de ellas incluyen la enseñanza de la pronunciación, la corrección fonética o la evaluación de la pronunciación, pero las posibilidades pueden ampliarse a los estudios que analizan los procesos de producción y percepción de una L2, al papel de los hablantes y de los oyentes, a los factores (lingüísticos y no lingüísticos) que intervienen en las reacciones de los interlocutores, o a las investigaciones que examinan las variables fónicas que contribuyen a ese acento y, dentro de él, a los elementos que específicamente impiden la comunicación (inteligibilidad) a los que la dificultan (comprensibilidad), o a los que solo son marca del acento extranjero. Considerar todos estos aspectos impone a su vez la necesidad de estudiar las dimensiones que configuran y definen el propio marco conceptual de «accento», ya que se constituye como requisito indispensable para establecer relaciones y campos de actuación más definidos y precisos. Con el propósito de arrojar algo de luz sobre estas dimensiones en el campo del español como L2, y de marcar un punto de partida desde el que avanzar hacia muchos de los anteriores aspectos, se ha realizado el presente estudio.

6. REFERENCIAS BIBLIOGRÁFICAS

- BARRERA PARDO, D. (2004): «Can pronunciation be taught? A review of research and implications for teaching», *Revista Alicantina de Estudios Ingleses* 17, pp. 31-44.
- BURDA, A.N., SCHERZ, J.A., HAGEMAN, C.F. y EDWARDS, H.T. (2003): «Age and understanding speakers with Spanish or Taiwanese accents», *Perceptual and Motor Skills* 97, 1, pp. 11-20.
- BUSCH, M. (1993): «Using Likert scales in L2 research. A researcher comments», *TESOL Quarterly* 27,4, pp. 733-736.
- CANTERO, F.J. (2003): «Fonética y didáctica de la pronunciación», en Mendoza, A. (ed.), *Didáctica de la lengua y la literatura*, Madrid, Prentice Hall, pp. 545-572.
- CHEN, H.C. (2011): «Judgments of intelligibility and foreign accent by listeners of different language backgrounds», *The Journal of Asia TEFL* 8, 4, pp. 61-83.
- DERWING, T.M. (2003): «What do ESL students say about their accents?», *Canadian modern language review/ La Revue Canadienne des Langues Vivantes* 59, 4, pp. 547-567.
- , (2008): «Curriculum issues in teaching pronunciation to second language learners», en Hansen Edwards, J. y Zampini, M. (eds), *Phonology and Second Language Acquisition*, Ámsterdam, John Benjamins, pp. 347-369.
- DERWING, T.M. y MUNRO, M.J. (1997): «Accent, intelligibility, and comprehensibility: Evidence from four L1s», *Studies in Second Language Acquisition* 20, 1, pp. 1-16.
- , (2009): «Putting accent in its place: rethinking obstacles to communication», *Language Teaching* 42, 4, pp. 476-490.
- DERWING, T.M., MUNRO, M.J. y THOMSON, R.I. (2007): «A longitudinal study of ESL learners' fluency and comprehensibility development», *Applied Linguistics* 29, 3, pp. 359-380.
- DERWING, T. y ROSSITER, M. (2004): «Second language fluency: Judgments on different tasks», *Language Learning* 54, 4, pp. 655-679.
- DERWING, T.M., ROSSITER, M. y MUNRO, M. (2002): «Teaching native speakers to listen to foreign-accented speech», *Journal of Multilingual and Multicultural Development* 32, 4, pp. 245-259.
- FAYER, J.M. y KRASINSKI, E. (1987): «Native and nonnative judgments of intelligibility and irritation», *Language Learning* 37, 3, pp. 313-326.
- FIELD, J. (2005): «Intelligibility and the listener: The role of lexical stress», *TESOL Quarterly* 39, 3, pp. 399-423.
- GALLARDO del PUERTO, F., GÓMEZ LACABEX, E. y GARCÍA LECUMBERRI, M.L. (2006): «The assessment of foreign accent by native and non-native judges», *Phonetics teaching and learning conference*, University College London, pp. 1-4.
- GALLEGO, J.C. (1990): «The intelligibility of three nonnative English-speaking teaching assistants: an analysis of student-reported communication breakdowns», *Issues in Applied Linguistics* 1, 2, pp. 219-237.
- GASS, S. y VARONIS, E. (1984): «The effect of familiarity on the comprehensibility of nonnative speech», *Language learning* 34, 1, pp. 65-89.
- GIL, J. (2007): *Fonética para profesores de español: de la teoría a la práctica*, Madrid, Arco/Libros.
- HAHN, L.D. (2004): «Primary stress and intelligibility: research to motivate the teaching of suprasegmentals», *TESOL Quarterly* 38, 2, pp. 201-223.

- INSTITUTO CERVANTES (2008): *Diccionario de términos clave de ELE*, Madrid, SGEL, Disponible también en versión electrónica en: <http://cvc.cervantes.es/ensenanza/biblioteca_ele/diccio_ele/diccionario/correccion.htm>.
- ISAACS, T. (2008): «Towards defining a valid assessment criterion of pronunciation proficiency in graduate students», *The Canadian Modern Language Review* 4, pp. 555-580.
- ISAACS, T. y TROFIMOVICH, P. (2012): «Deconstructing comprehensibility», *Studies in Second Language Acquisition* 34, 3, pp. 475-505.
- JENKINS, J. (2000): *The phonology of English as an international language: New models, new norms, new goals*, Oxford, Oxford University Press.
- KANG, O. (2010): «Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness», *System* 38, 2, pp. 301-315.
- KANG, O. y RUBIN, D.L. (2009): «Reverse linguistic stereotyping: measuring the effect of listener expectations on speech evaluation», *Journal of Language and Social Psychology* 28, 4, pp. 441-456.
- KENNEDY, S. y TROFIMOVICH, P. (2008): «Intelligibility, comprehensibility, and accentedness of L2 speech: The role of listener experience and semantic context», *The Canadian Modern Language Review* 64, 3, pp. 459-489.
- KORMOS, J. y DÉNES, M. (2004): «Exploring measures and perceptions of fluency in the speech of second language learners», *System* 32, 2, pp. 145-164.
- LENNON, P. (1990): «Investigating fluency in EFL: A quantitative approach», *Language Learning* 40, 3, pp. 387-417.
- LEVIS, J.M. (2005): «Changing contexts and shifting paradigms in pronunciation teaching», *TESOL Quarterly* 39,3, pp. 369-377.
- LIPPI-GREEN, R. (2004): «Language Ideology and language prejudice», en Finegan, E. y Rickford, J. (eds), *Language in the USA: Themes for the Twenty-First Century*, Cambridge, Cambridge University Press, pp. 289-304.
- LLISTERRI, J. (2003): «La evaluación de la pronunciación en la enseñanza del español como segunda lengua», en Reyzabal, M.V. (ed.), *Perspectivas teóricas y metodológicas: Lengua de acogida, educación intercultural y contextos inclusivos*, Consejería de Educación, Comunidad de Madrid, pp. 547-562.
- LLURDA, E. (1995): «Native-speaker reactions to nonnative speech: A review», *Sintagma* 7, pp. 43-51.
- LUDWIG, J. (1982): «Native-speaker judgments of Second-Language learners' efforts at communication: A review», *The Modern Language Journal* 66, 3, pp. 274-283.
- MUNRO, M.J. (2003): «A primer on accent discrimination in the Canadian context», *TESL Canada Journal* 20, 2, pp. 38-51.
- MUNRO, M.J. y DERWING, T.M. (1994): «Evaluations of foreign accent in extemporaneous and read material», *Language Testing* 11, 3, pp. 253-266.
- , (1995a): «Foreign accent, comprehensibility, and intelligibility in the speech of second language learners», *Language Learning* 45, 1, pp. 73-97.
- 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, pp. 289-306.
- , (1999): «Foreign accent, comprehensibility, and Intelligibility in the speech of second language learners», *Language Learning* 49, pp. 285-310.

- ODISHO, E.Y. (2003): *Techniques of teaching pronunciation in ESL, bilingual and foreign language classes*, Múnich, Lincom Europa.
- PISKE, T., MACKAY, I. y FLEGE, J. (2001): «Factors affecting degree of foreign accent in an L2: A review», *Journal of Phonetics* 29, 2, pp. 191-215.
- POWELL CROCKER, L. (2010): *Intelligibility of Hindi-accented English: the role of duration and intonation*, University of Minnesota.
- RAJADURAI, J. (2007): «Intelligibility studies: A consideration of empirical and ideological issues», *World Englishes* 26, 1, pp. 87-98.
- RAUX, A. y KAWAHARA, T. (2002): «Automatic intelligibility assessment and diagnosis of critical pronunciation errors for computer-assisted pronunciation learning», *ICSLP* 1, pp. 737-740.
- RINEY, T.J. (2005): «Phonetic parameters and perceptual judgments of accent in English by American and Japanese listeners», *TESOL Quarterly* 39, 3, pp. 441-466.
- SANTAMARÍA BUSTO, E. (2013): «Enseñar la competencia fonética», en Ruiz de Zarobe, L. y Ruiz de Zarobe Y. (eds.), *Enseñar hoy una lengua extranjera*, Berkeley, Portal Editions, pp. 2-66.
- SCHAIRER, K.E. (1992): «Native speaker reaction to non native speech», *The Modern Language Journal* 76, 3, pp. 309-319.
- SINGLETON, D. y RYAN L. (2004): *Language acquisition: The age factor*, Clevedon, Multilingual Matters.
- SMITH, L.E. y NELSON, C.L. (1985): «International intelligibility of English: directions and resources», *World Englishes* 4, 3, pp. 333-342.
- TROFIMOVICH, P. y ISAACS, T. (2012): «Disentangling accent from comprehensibility», *Bilingualism: Language and Cognition* 15, 4, pp. 905-916.
- WINTERS, S. y O'BRIEN, M.G. (2013): «Perceived accentedness and intelligibility: The relative contributions of F0 and duration», *Speech Communication* 55, 3, pp. 486-507.

Edita
SeL

