



Comunicación y Sociedad
ISSN: 0188-252X
comysoc@yahoo.com.mx
Universidad de Guadalajara
México

Soto Sanfiel, María Teresa
Efecto del tono de voz y de la percepción del rostro en la formación de impresiones sobre los
hablantes mediáticos
Comunicación y Sociedad, núm. 10, julio-diciembre, 2008, pp. 129-161
Universidad de Guadalajara
Zapopan, México

Disponible en: <http://www.redalyc.org/articulo.oa?id=34601006>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica
Red de Revistas Científicas de América Latina, el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

Efecto del tono de voz y de la percepción del rostro en la formación de impresiones sobre los hablantes mediáticos

MARÍA TERESA SOTO SANFIEL¹

Se observó el efecto de la frecuencia fundamental de las voces (el tono) y de la percepción del rostro en las impresiones sobre hablantes mediáticos. Se pretendía determinar si existía relación entre la información visual y una característica eminentemente acústica: el tono. 320 sujetos fueron sometidos a una de dos condiciones: sonora (percibían sólo las voces) o audiovisual (veían las caras). Luego, respondían mediante escalas del diferencial, a la pregunta: ¿El locutor le pareció? Los resultados muestran que el tono influye en las impresiones, que la percepción del rostro evidencia las diferencias en el enjuiciamiento de las voces agudas, medias o graves, y que la credibilidad dirige las evaluaciones.

PALABRAS CLAVE: interpretación, voz, rostro, formación de impresiones, efectos.

The study observed the effect of fundamental frequency of voices and perception of the face on the audience's impressions of radio and television speakers. It pretended to determine if there was a relationship between the visual information obtained from face and a specific attribute of oral interpretation: tone. 320 subjects were assigned to one of two conditions: audio (only hearing voices) or audiovisual (also seeing faces). Later, answered a questionnaire with semantic differential scales. Results show tone does influence impressions, that perception of face does make a difference to judgments depending of the tone of the voices (low, medium or high) and that credibility directs evaluations of media speakers.

KEY WORDS: interpretation, voice, face, formation of impressions, effects.

¹ Universidad Autónoma de Barcelona.

Correo electrónico: Mariateresa.soto@uab.es

INTRODUCCIÓN

El objetivo general de este trabajo es contribuir a la obtención de conocimientos sobre los procesos de formación de impresiones en contextos de comunicación masiva y avanzar en el estudio sobre el enjuiciamiento de personas que hablan a través de los medios de comunicación de difusión electrónica (en contexto de habla hispana, además). Más específicamente, aunque todavía de modo amplio, este trabajo pretende obtener conocimientos sobre el efecto que, en la formación de impresiones de los receptores sobre los hablantes mediáticos, tiene la disponibilidad de determinados canales de comunicación no verbal –la voz o el rostro– de los locutores (como sucede en radio o en televisión). En este sentido, es cierto que existen numerosos estudios del área de la comunicación interpersonal que afirman que tanto las expresiones faciales como el tono de la voz son instrumentos altamente informativos y enormemente potentes para la comunicación de estados internos, actitudes o sentimientos (De Paulo y Friedman, 1998:3-39) y que, de hecho, pueden informar a los perceptores de las habilidades y personalidad de los otros (Ambady, Bernieri y Richeson, 2000:201-271). Sin embargo, existe insuficiencia de estudios que exploren estos fenómenos en la comunicación mediada.

Ahora bien, específicamente, este texto tiene por objetivo dar detallada cuenta de una investigación experimental diseñada para estudiar el efecto de la modalidad de percepción (y con ello de la disponibilidad de claves visuales y sonoras) en la formación de impresiones sobre los locutores mediáticos (los que nos hablan –sea a través de la radio o de la televisión– para transmitir mensajes complejos, como una noticia o un monólogo de ficción). La investigación que se relata sopesa el efecto que tienen la voz y el rostro de los locutores al preguntarse cómo se relacionan ambos tipos de señales en la formación de impresiones de las audiencias: ¿existen diferencias en nuestras opiniones sobre un locutor cuando le oímos únicamente (como en radio) o cuando también le vemos (como en televisión)? Esta es la pregunta para cuya respuesta se diseñó la investigación que se presenta.

Asimismo, además de explorar si existe un efecto de la percepción del rostro –y, por lo tanto, de la modalidad de percepción– en la formación de impresiones sobre los locutores, el diseño experimental que se

aplicó buscaba profundizar en una propiedad particular de las voces. De hecho, buscaba determinar si en la percepción de los hablantes existía una relación entre la información visual del rostro que producía el habla y una característica eminentemente acústica de la interpretación: el tono o frecuencia fundamental de la voz. Luego, este texto ofrece los resultados de un experimento que sopesó dos tipos de señales no verbales que se producen simultáneamente durante la percepción y producción del habla (las visuales y las sonoras): ¿existe un efecto distinto en la influencia que tiene la percepción del rostro dependiendo de si los hablantes tienen tonos de voces graves, medios o agudos? ¿Afecta la percepción del rostro de distinta o igual manera al enjuiciamiento de la personalidad de las voces con distintos tonos? ¿A cuál locutor radiofónico influiría más su aparición en televisión: al que tiene la voz aguda o media o al que tiene la voz grave? El cómo se llegó a formular preguntas tan específicas se relata en el punto que sigue a continuación. Sin embargo, antes de proseguir, explicaremos la estructura de este texto.

Advertimos, en principio, que este trabajo es el reporte de un trabajo que aplica el método experimental, que es el que permite establecer relaciones causales entre distintas variables. Mediante la aplicación de este método, se producen situaciones artificiales, de laboratorio, para probar la existencia de un efecto de una (o varias) variables independientes (causas) sobre otras dependientes (efecto). El reporte de los resultados de un experimento implica dar cuenta completa del método para favorecer su replicabilidad, esto es, para que cualquier otro investigador pueda repetir el trabajo y llegar a conclusiones semejantes. El lector, entonces, encontrará la información esencial de cada paso del desarrollo.

Aparte de esta introducción, el artículo está organizado en cinco partes. En la primera se ofrecen las evidencias sobre las que descansa la formulación de las hipótesis de la investigación (las que, a su vez, se incluyen cuando ya se ha presentado la información teórica que las justifica). En la segunda parte se expone el método. En ella se explica el diseño experimental, los materiales usados en las pruebas de percepción y las variables dependientes (las medidas por los sujetos) que se aplicaron. En la tercera parte se cuentan tanto las pruebas como los procedimientos que se siguieron para el análisis estadístico de los datos obtenidos de las pruebas de percepción. En la siguiente, se exponen

los resultados obtenidos. Finalmente, en la quinta parte, se ofrecen las conclusiones relevantes.

La autora agradece a los doctores Emili Prado, José Reinaldo Martínez y José Ignacio Latorre su colaboración en las distintas fases del estudio.

ANTECEDENTES

La formación de impresiones: datos básicos

Los procesos por los que los seres humanos combinamos las informaciones de que disponemos para generar opiniones globales acerca de otros individuos, se conocen con el nombre de formación de impresiones. *Grosso modo*, la literatura define a la formación de impresiones como los intentos que realizamos para que los datos que obtenemos sobre otras personas se ajusten a unas categorías que hemos adquirido de antemano (Fiske y Neuberg, 1990:1-73). Se acepta, en general, que los humanos hemos desarrollado estos mecanismos de categorización del otro para compensar nuestra limitada capacidad de procesamiento de información novedosa (Abele y Petzold, 1998:347-358).

El procesamiento –la formación de impresiones– comienza cuando de manera espontánea, inclusive sin darnos cuenta, asociamos rasgos de la personalidad del individuo que percibimos con características acumuladas en nuestra cognición (Newman y Uleman, 1990:224-240; Uleman y Newman, 1992:77-90). El proceso supone el manejo de ingentes paquetes de información sobre lo observado, debido a que extraemos rasgos detallados del comportamiento, o de la apariencia, del otro y los relacionamos activamente para dotarles de significado. Dada la gran cantidad de datos que empleamos en la categorización, tendemos a hacer perdurar las valoraciones que nos ayudan a explicar a quienes contemplamos. De hecho, estas son muy difíciles de cambiar (Fiske y Taylor, 1991). Los investigadores de lo social estudian con profusión los fenómenos relacionados con la formación de impresiones, ya que ellos, por su presencia constante, explican, en parte, las dinámicas de la interacción social.

Ahora bien, la formación de impresiones ha sido estudiada, preferentemente, en contextos de comunicación cara a cara y los conocimientos acumulados sobre el tema se aplican, con suma frecuencia en la literatura

académica, a la explicación de fenómenos ocurridos en ámbitos de la comunicación mediada. Así, en ocasiones se da por cierto que la formación de impresiones sobre los actores, los profesionales de la representación, aquellos que se dirigen a las audiencias, a través de la radio o la televisión para transmitir mensajes complejos, se puede explicar satisfactoriamente aplicando sin discriminación los conocimientos obtenidos sobre el proceso de formación de impresiones en comunicación interpersonal. Sin embargo, se dispone de pruebas a favor de la existencia de diferencias entre los procesos de ambos contextos comunicativos (Jacobson, 1999; Walther, 1993:381-389, 1996:3-43). Parece que, como manifiesta Walter (1993), las diferencias en las evaluaciones sobre los individuos, según la modalidad en que se perciben, se deben a que la mediación tiende a inhibir el efecto de las claves no verbales en el proceso del enjuiciamiento del otro, y ello produce, evidentemente, resultados perceptivos distintos a los de la interacción personalizada. No obstante esta evidencia, y a pesar de que la comunicación mediada es parte fundamental de la vida de los ciudadanos de las sociedades actuales, el estudio de la formación de impresiones, en estos contextos, demanda mayor atención (Sherman *et al.*, 2001:123-129; Collins Tidwell y Walter, 2002:317-348; Weisband, Schneider y Connolly, 1995:1124-1151). En particular, se echan de menos estudios que exploren esos procesos en el terreno concreto de los medios de comunicación audiovisual (*vgr.* radio o televisión).

La voz como fuente de información sobre el hablante

Que la voz afecta a los procesos de formación de impresiones es algo ampliamente reconocido por los investigadores de la comunicación interpersonal, aunque son menos los trabajos que exploren las propiedades de la voz en el enjuiciamiento de la personalidad de los hablantes mediáticos. Acercándonos a nuestro contexto, el de la comunicación electrónica de difusión masiva, hallamos el trabajo de Zuckerman y Miyake (1993) que pone de relieve la importancia del estudio del atractivo vocal en comunicación y justifica, asimismo, buena parte del presente estudio.

Estos investigadores probaron, en primer término, que los juicios sobre la personalidad de los hablantes –la formación de impresiones– se relacionaban con la percepción de atractivo de las voces. Según los resultados del trabajo, la magnitud del efecto de la variable “atractivo

vocal” en los juicios sobre la personalidad era realmente considerable. A pesar de ello, como demostraron, los receptores no eran conscientes de que percibían el atractivo en las voces y de que, espontáneamente, usaban esa información para evaluar la personalidad del otro.

Aparte de sugerir la gran importancia que tiene una voz atractiva en comunicación y en la formación de impresiones, el trabajo de Zuckerman y Miyake también mostró que los seres humanos asociábamos las voces atractivas con impresiones más positivas de la personalidad global del hablante; que la percepción de voces atractivas estimulaban nuestro deseo por afiliación, aumentaban el índice de similaridad que asumíamos con respecto al hablante y, además, permitían que percibiéramos a quien hablaba como de un estatus semejante al nuestro.

Zuckerman y Miyake, sin embargo, llegaron un poco más lejos y probaron, también, que existía una relación entre el atractivo vocal y medidas objetivamente obtenidas como la frecuencia fundamental, la amplitud y la duración del habla. El estudio, específicamente, halló una relación entre el tono de la voz y el atractivo percibido. Las voces con frecuencias fundamentales más bajas eran consideradas más atractivas. De hecho, los receptores asociaban a los tonos bajos con rasgos más favorables de la personalidad.

Esta última evidencia fue confirmada por un trabajo posterior. En 2000, Collins también exploró la relación entre la frecuencia fundamental y el atractivo vocal y halló que, de acuerdo con Zuckerman y Miyake, las voces masculinas con frecuencias más bajas eran consideradas más atractivas por los receptores femeninos. Collins, por otra parte, profundizó en los aspectos físicos que se percibían a partir de las voces. De ello hablamos a continuación.

La investigadora probó que los receptores deducían características del físico de los hablantes a partir de la percepción de sus voces y que ello sucedía aunque no se vieran las caras. También, reportó que el grado de coincidencia en la atribución de atractivo físico a los hablantes era muy alto entre los receptores. Profundizando, encontró que el peso de los hablantes era una de las características inferidas con mayor precisión a partir de datos objetivos (*vgr.* la frecuencia fundamental), mientras que otros atributos (*vgr.* la edad, el peso o la altura del hablante) no eran inferidos, de manera tan coincidente, entre los juicios.

Llegados a este punto, es preciso recordar que, aparte de la investigación de Zuckerman y Miyake, o la de Collins, otros estudios sobre formación de impresiones relacionan el tono de la voz con factores de la personalidad del hablante, como la madurez (Montepare y Zebrowitz-McCarthy, 1989:189-203), la edad (Caruso, Mueller y Shadden, 1995:63-80), la benevolencia (Riding, Lonsdale, Brown, 2006), la sensibilidad no verbal (Bond *et al.*, 1987:335-380) o con índices de estrés, emociones positivas, inestabilidad emocional y/o tensión psicológica (vgr. Apple, Streeter y Krauss, 1979; Fairbanks, 1940; Fairbanks y Pronovost, 1939; Laukka, Juslin y Bresin, 2005; Scherer, 1979:147-209 y 1978).

Junto a las evidencias aportadas coexiste un trabajo que es un referente inmediato del presente porque explora la formación de impresiones a partir de la voz en un entorno comunicativo radiofónico. El estudio al que nos referiremos a continuación, además de ser próximo por su orientación mediática, es cercano geográfica e idiomáticamente al contexto del experimento reportado en este artículo. Nótese, asimismo, que es anterior al de Zuckerman y Miyake y al de Collins.

Para comenzar, Rodríguez (1989) dividió a los rasgos de las voces que transmiten información sobre las imágenes de los emisores en cuatro niveles, según el tipo de información transmitida: “idiográfico” (relativo al aspecto físico de los emisores), “caracterial o afectivo” (relacionado con el carácter o la actitud emocional), “encuadrático” (vinculado al entorno social) y “sintomático” (que se refiere a trastornos físicos o psíquicos). Su estudio experimental se centró en los elementos sonoros vinculados al aspecto físico y al carácter o actitud de las voces radiofónicas.

El investigador halló que, efectivamente, las audiencias eran capaces de formarse una imagen física y psicológica de los hablantes en un grado de coincidencia superior al definido por el azar (lo que es coherente con las pruebas que Collins aportara tiempo después). Además, reportó que esas imágenes formadas en la mente de las audiencias coincidían con la imagen real de los locutores.

Rodríguez profundizó un poco más y halló que la percepción de agrado en las voces era el factor más influyente en la determinación de “radiogenia”; que una voz *radiogénica* (la ideal para radio) era aquella percibida como agradable. También, encontró que el agrado se asociaba con el

atractivo percibido en el hablante. Según Rodríguez, las voces agradables creaban en la mente del oyente, de forma sistemática, una imagen atractiva y bella. A pesar de todo esto, el estudio falló en encontrar rasgos físicos concretos que se pudieran asociar a la llamada radiogenia.

En una tercera fase, finalmente, realizada mediante análisis espectrales, el investigador halló una relación entre la inteligencia percibida en los hablantes y el tono de la voz (la frecuencia de las voces). Según los resultados aportados, una voz percibida como inteligente tenía resonancias agudas pero mantenía su intensidad general concentrando sus pocos descensos sólo en determinados formantes del espectro (F0 y F3). Parecía, por lo tanto, que el tono estaba relacionado con el atractivo percibido y con algún factor que explicaba la radiogenia de los hablantes radiofónicos.

Todos los trabajos citados avalan que, tal y como categorizó Scherer (1979), los seres humanos obtenemos dos tipos de información del habla. Por una parte, adquirimos datos de los aspectos del contenido (*vgr.* las palabras o las frases) y, por otra, de las cualidades funcionales de la voz, esto es, de los parámetros acústicos (*vgr.* tono, timbre, velocidad del habla e intensidad). El trabajo que se reporta en este artículo se centra, específicamente, en la exploración del efecto de una variable –la percepción del rostro del hablante–, en una cualidad funcional específica del habla: el tono o frecuencia fundamental, y, por tanto, explora parte del segundo tipo de informaciones categorizadas por Scherer. Sin embargo, como paso previo, este trabajo también observa el enjuiciamiento de las audiencias a los locutores, según sea el tono de su voz (agudo, medio y grave), independientemente de si ve o no su rostro (de la modalidad de percepción). El objetivo de este primer análisis, como se puede deducir, es comprobar, tal y como previene la literatura, si existe un efecto de la frecuencia fundamental de las voces de los hablantes mediáticos en la formación de impresiones de sus audiencias, partiendo de las siguientes hipótesis:

H1: La frecuencia fundamental de las voces de los locutores influye en la formación de impresiones sobre los hablantes.

Ahora bien, hemos advertido que este trabajo tenía también por objetivo observar el efecto de la percepción del rostro en la formación de impresiones según fuese la frecuencia fundamental de la voz de los locutores. Por ello, antes de proseguir, es prudente justificar por qué

hemos creído que podría haber una relación en la evaluación de los receptores y las variables frecuencia fundamental-percepción del rostro. Esto es, defender por qué se cree, en la investigación que se reporta, que la percepción del rostro influye de distinta manera según se enjuicie a un locutor de voz aguda, media o grave.

*De la influencia de la percepción del rostro en el enjuiciamiento
(el habla audiovisual)*

Para hipotetizar sobre la relación entre la voz y el rostro en la formación de impresiones se debe tomar primeramente en cuenta a los conocimientos actuales sobre percepción del habla bimodal. Estas pruebas, procedentes de la psicoacústica, y observadas a la luz de la comunicación mediática, permiten sostener que existen diferencias entre escuchar a un hablante por radio o por televisión, ya desde el mismo procesamiento perceptivo de las señales que recibe el receptor.

Así, los estudios sobre percepción del habla audiovisual muestran que los seres humanos integramos perceptivamente las señales sonoras y visuales de los eventos del habla, y que dicha integración produce un resultado perceptivo distinto al que produce el procesamiento de cada canal por separado. La demostración de la integración bimodal del habla se conoce como “efecto o ilusión McGurk” (MacDonald y McGurk, 1978:253-257), y confirma que a un nivel de procesamiento aún desconocido se produce una sinergia, una interacción de la información provista por ambos canales. El fenómeno prueba que si la sílaba auditiva *ba* se muestra junto a un video del hablante que dice la sílaba *ga*, los receptores perciben la sílaba *da*; una sílaba nueva e inexistente en las modalidades sonora y visual por separado. La percepción de los movimientos articulatorios del hablante, por lo tanto, modifica significativamente la experiencia perceptiva de la escucha.

La demostración de la existencia del efecto McGurk provocó un interés por estudiar las propiedades del proceso de percepción del habla bimodal. Por eso, se ha demostrado que el resultado perceptivo del habla audiovisual es una unidad de creencia sólida formada por la combinación de los canales auditivo y visual que se resiste a la identificación fragmentada de sus componentes parciales (Massaro y Cohen, 1983:753-771; Summerfield, 1987:3-51); que los receptores extraen

y utilizan los datos que cada canal proporciona, por lo que obtienen informaciones distintas de cada canal (Massaro y Cohen, 1983); que la integración perceptiva es el resultado de un proceso en el que los receptores evalúan las informaciones, otorgan grados de soporte a las opciones y toman una decisión; que los receptores utilizan las fuentes de información auditiva y visual para distintos propósitos y que el proceso produce el “realce” de uno de los dos canales, más que una solución intermedia (Summerfield, 1987). Finalmente, se ha probado también que, en el procesamiento del habla bimodal, la señal audible parece ser más influyente que la visible (Massaro y Cohen, 1996:753-771).

No obstante, a las anteriores evidencias se añade el trabajo descrito por Strand (1998). Los resultados reportados por esta investigadora dicen que la percepción audiovisual del hablante supone, efectivamente, la integración de la información visual con la sonora pero que, también, envuelve el acceso a expectativas (estereotipos) de los receptores acerca de cómo debería sonar el habla en función de cómo lucen los hablantes. Las evidencias que obtuvo en su trabajo con E. Johnson muestran, según Strand, que los receptores tienden a percibir los límites fonológicos a frecuencias del espectro menores si están acompañadas de la cara de un hombre y a mayores si la cara era de una mujer. Es decir, el tono de la voz se percibe más bajo si la cara que lo acompañaba es masculina y más alta si es femenina. Por ello, y especialmente relacionado con el trabajo que aquí se presenta, el estudio de Strand demuestra que las expectativas sociales complejas tienen una influencia en cómo se categoriza la señal sonora del habla audiovisual.

Otros trabajos, en línea semejante (Niedzielski, 1999:62-85; Daly y Bench, 1996:468-480; Wyer *et al.*, 1994:254-267), confirman que, en la percepción del habla audiovisual, los humanos usamos tanto la información social como la visual para calibrar o crear el espacio fonológico de los hablantes. Específicamente, Wyer *et al.* probaron que las claves visuales invitaban a los receptores a hacerse inferencias generales, a pensar acerca de características globales –estereotípicas– más que en atributos específicos individuales en los procesos de formación de impresiones del habla bimodal. Según estos investigadores, las claves visuales se relacionaban con comportamientos estereotípicos esperados por el oyente en el hablante, según fuese el contexto comunicativo, y

esto último, a su vez, afectaba a la formación de impresiones. Por su parte, el trabajo de Daly y Bench señala que la disponibilidad de la información visible del habla (los movimientos articulatorios) afecta al enjuiciamiento del hablante porque los receptores recurren a la información visual para dar peso a un elemento que no tiene fuerza significativa en el enjuiciamiento a partir de la voz, pero que sí la tiene desde la percepción del rostro. Los investigadores observaron que los participantes de su experimento emitían juicios semejantes después de ver el discurso (sin audio) y después de oírlo. En ambas situaciones experimentales, las variables *status* (*educated*, *reliable* y *competent*), en primer lugar, y *solidarity* (*pleasant*, *friendly*, *attractive* y *trustworthy*) en segundo, aparecían relacionadas con los juicios. Sin embargo, únicamente con la percepción del lenguaje visual aparecía en el conjunto un tercer factor –*performance* (*expressivity*, *natural*, *relaxed*)–, que afectaba estadísticamente a la formación de impresiones. Los tres factores, conjuntamente, explicaban 65% de la varianza. Por sí mismos, *status* explicaba 42%, *solidarity*, 14% y *performance* 8%. Por lo anterior, asimismo, Daly y Bench concluían que, en la formación de impresiones, la información sonora soportaba un mayor peso en el enjuiciamiento de los hablantes que la información visual.

Las pruebas aportadas conducen a pensar que la información visual –obtenida durante el procesamiento bimodal– es funcional y pragmáticamente orientada, por los receptores, a la formación de impresiones, y que la señal audible parece ser más influyente que la visible en el enjuiciamiento de los hablantes. Sin embargo, y a pesar de ello, se sabe que los individuos valoramos positivamente la disponibilidad de mayor información sobre las personas a las que enjuiciamos porque se acrecienta nuestra confianza en el enjuiciamiento. También parece que dicha confianza no tiende, en modo alguno, a aumentar la precisión de nuestros juicios (Dunning *et al.*, 1990:568-581).

Finalmente, debe apuntarse que, tal y como se ha avanzado, desconocemos la existencia de trabajos que exploren la formación de impresiones a la luz de los medios de comunicación social y de las teorías que los explican. Sin embargo, disponemos de la “teoría de la riqueza de los medios” (*richness media theory*) que explica (y categoriza) modelos de procesamiento de la información. La teoría cree que el éxito en la

consecución de los objetivos de las organizaciones reside, en gran parte, en la mayor o menor disponibilidad de información que procuran a sus públicos, especialmente en la cantidad de claves visuales que ofrecen (Daft y Lengel, 1984:191-233). La teoría postula que los medios más ricos facilitan la reducción de imprecisiones porque permiten compartir significados (Treviño, Lengel y Daft, 1987:553-575). Aplicada a los medios de comunicación, esta teoría permite suponer, por ejemplo, que la televisión, a diferencia de la radio, podría favorecer la homogeneidad en la formación de impresiones de las audiencias dado que ofrece mayor información (claves visuales) sobre los fenómenos.

Por todo lo anterior, este estudio, finalmente, también hipotetiza que H2: la percepción audiovisual del rostro del hablante afectará a las evaluaciones de los sujetos de manera diferente según sea la frecuencia fundamental de las voces de los locutores.

MÉTODO

Diseño experimental

Los participantes fueron 320 estudiantes de ambos géneros, de una facultad española, con una edad promedio de 21.03 años ($SD = 2.178$).

Los sujetos fueron asignados equitativa y aleatoriamente a una de dos condiciones experimentales: sonora o audiovisual. El *corpus* fue el mismo para ambas condiciones, de manera que las variaciones en el enjuiciamiento se debieran, únicamente, a la modalidad de percepción. En la condición sonora los sujetos escuchaban las voces; en la audiovisual, además de escucharles, veían las imágenes de los locutores mientras hablaban. Todas las pruebas se hicieron en una misma aula.

Los materiales fueron transmitidos a través de un televisor con pantalla de 24 pulgadas. En las pruebas sonoras la pantalla aparecía vacía porque se reducían completamente los niveles de brillo y contraste del aparato; en las audiovisuales, los parámetros se colocaban a sus niveles normales. Así, se aseguró una de las principales condiciones de la situación experimental: que los niveles acústicos de la percepción de las voces de los locutores fuesen exactamente los mismos en todas las pruebas y condiciones.

A los sujetos se les informaba que el estudio tenía por objetivo evaluar a locutores mediáticos. La identificación de los investigadores que aplicaban las pruebas como pertenecientes a la Facultad de Comunicación Audiovisual, y la etiqueta “locutores” con que, en todas las instrucciones del experimento, se identificaba a los hablantes, pretendían circunscribir el enjuiciamiento al contexto de la comunicación audiovisual. Existen investigaciones que certifican que los receptores son susceptibles a las instrucciones que reciben en los experimentos hasta el punto de que esa susceptibilidad influye en la selección de los factores que les permiten formarse impresiones y categorizar los comportamientos que perciben. De hecho, los sujetos experimentales son capaces de usar las etiquetas con que se presentan los estímulos, de forma dinámica y orientada a la interpretación de los comportamientos que observan, incluso de aquellos ambiguos (Higgins, Rholes y Jones, 1977:141-154).

El investigador informaba a los sujetos que percibirían las interpretaciones de ocho locutores y de que cada interpretación duraba cerca de un minuto. También que, tras la recepción de cada hablante, detendría la emisión para que pudieran contestar el cuestionario que contenía escalas del diferencial semántico, precedidas de la pregunta: “¿El locutor le pareció?”

Materiales

Se diseñaron dos textos que serían interpretados por los locutores. Se acordó que aquellos fuesen distintos para neutralizar el efecto del contenido sobre las impresiones y provocar una escucha variada. Asimismo, se convino que cada texto durase un minuto. Por otra parte, se decidió que el tema de estos debía ser inocuo para evitar reacciones negativas que pudiesen afectar a la credibilidad de los locutores. Uno de los textos era una noticia y el otro un cuento.

Se solicitó la colaboración a ocho locutores femeninos y ocho masculinos, a quienes se había escogido tomando en consideración que tuvieran tonos de voz agudos, medios y graves, además de un habla con vocalización clara e inteligible. Los locutores no eran conocidos ni populares.

Los materiales se grabaron en video, en un estudio de radio que tenía las condiciones para garantizar la calidad audiovisual profesional de los materiales. La cámara se situó en la sala de control del estudio y captó a

los locutores, que estaban en la cabina, a través de los cristales del locutorio. Se utilizó únicamente la luz artificial del estudio porque era suficiente para capturar adecuadamente el rostro completo de los hablantes.

La cámara captó un plano medio corto frontal de los locutores que permitía ver sus rostros e interpretaciones visuales completos. Debido a que la frecuencia fundamental era una variable observada, no se manipuló ni ecualizó ninguno de los parámetros acústicos. Se determinó una intensidad de captación promedio de las voces, de antemano, que se mantuvo estándar en todas las grabaciones.

Estas se realizaron con un micrófono profesional (MILAB Sweeden VIP-50), conectado a una mezcladora profesional de sonido (TASCAM M-3500) que, a su vez, estaba conectada a una cámara de video (SONY DXC-327P). Esta tenía lentes ópticos profesionales (FUJINON 1:1.4/7.5-90 mm) y grabador (EVV-9000P)

A los locutores se les informó que la investigación tenía por objetivo estudiar los factores acústicos de las voces y que serían grabados en video para que pudieran ver sus rostros durante la producción del habla. También, que debían interpretar dos textos distintos, de alrededor de un minuto de duración cada uno, y que tenían libertad para escoger el modo de representación de cada texto. Previamente, los textos habían sido impresos en un cartel con un tamaño de letra que permitía que se pudiera leer con facilidad desde el emplazamiento del micrófono y los textos habían sido colocados en una pared de la cabina, a manera de *telemprompter*. Los investigadores tomaron en cuenta que su lectura no provocara el descenso de la mirada del locutor.

Los locutores pudieron ensayar los textos el tiempo que consideraron necesario. Se les había informado que su actuación no podía tener errores, dudas, silencios injustificados o palabras cortadas y que si se equivocaban en la elocución podían repetirla. El ensayo de los locutores se justifica por la existencia de suficientes evidencias a favor de una relación entre habla fluida, dudas, repeticiones o frases interrumpidas y las evaluaciones de los receptores (Barge, Schluter y Pritchard, 1989; Burgoon, 1978; Burgoon, Birk y Pfau, 1990; Ericsson, Lind, Jonson y O'Barr, 1978; McCroskey y Mehrley, 1969; Miller y Hewegill, 1964; Ostermeier, 1967; Scherer, London y Wolf, 1973; Sereno y Hawkins, 1967).

Una vez que se habían realizado las grabaciones, se procedió a obtener la frecuencia fundamental promedio de cada voz con ayuda del programa *Mac Speech Lab*, que permite la medición espectral de formas electroacústicas. Se calculó un promedio general de la frecuencia fundamental de las voces que participaron en el experimento, por sexo, para crear las propiedades de la variable: aguda, media y grave. Para los hablantes femeninos se estableció que la frecuencia aguda comprendería las voces entre 189-225 Hz; la media tendría a las voces entre 152-188 Hz y la grave acogería a las voces entre 115-151 Hz. Para los hablantes masculinos se determinó que la aguda se referiría a las voces entre 152-178 Hz; la media a las voces entre 126-151 Hz y la grave a las voces entre 98-125 Hz.

De los locutores que participaron en las grabaciones, los investigadores seleccionaron a 16 que formarían parte del corpus experimental con el siguiente criterio: diferentes frecuencias fundamentales de las voces de cada sexo e interpretación acústica y visual variada sin errores (en el contenido verbal y en el comportamiento visual). Se consideró, también, que su actuación visual no pareciera nerviosa. Los seleccionados fueron dos hablantes de cada sexo con frecuencias agudas, cuatro hablantes de cada sexo con frecuencias medias y dos de cada sexo con frecuencias graves.

La *posproducción* del “corpus experimental”, la edición de las secuencias de visionado, se hizo tomando en cuenta que todos los locutores pudiesen ser enjuiciados mientras interpretaban los dos tipos de textos. Un ejercicio de combinatoria permitió crear ocho grupos de percepción, tomando en consideración que los locutores aparecieran sólo una vez en cada grupo, interpretando solo un texto. Asimismo, la distribución de los locutores en los grupos de percepción persiguió que no hubiese diferencias muy notorias en las frecuencias de las voces. Para atenuar las diferencias, se alternaron los locutores de ambos sexos (véase cuadro 1).

Variables dependientes

Para el análisis se usó un conjunto de indicadores asociados a la formación de impresiones sobre hablantes radiofónicos reportada por trabajos previos (Prado, 1992, 1997; Soto, 2000): agrado, credibilidad, familia-

CUADRO 1
DISTRIBUCIÓN DE LOS LOCUTORES EN LOS GRUPOS DE PERCEPCIÓN

P	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5	Grupo 6	Grupo 7	Grupo 8
1	M1 - FFA - T1	M4 - FFM - T1	M3 - FFG - T2	M2 - FFA - T2	F5 - FFM - T2	F8 - FFA - T2	F7 - FFG T1	F6 - FFM - T1
2	F1 - FFA - T2	F4 - FFG - T2	F3 - FFG - T1	F2 - FFG - T1	M5 - FFM - T1	M8 - FFA - T1	M7 - FFM - T2	M6 - FFM - T2
3	M2 - FFA - T1	M1 - FFA - T1	M4 - FFM - T2	M3 - FFG - T2	F6 - FFM - T2	F5 - FFM - T2	F8 - FFA - T1	F7 - FFG - T1
4	F2 - FFG - T2	F1 - FFA - T2	F4 - FFG - T1	F3 - FFG - T1	M6 - FFG - T1	M5 - FFM - T1	M8 - FFA - T2	M7 - FFM - T2
5	M3 - FFG - T1	M2 - FFA - T1	M1 - FFA - T2	M4 - FFM - T2	F7 - FFG - T2	F6 - FFM - T2	F5 - FFM - T1	F8 - FFA - T1
6	F3 - FFG - T2	F2 - FFG - T2	F1 - FFA - T1	F4 - FFG - T1	M7 - FFM - T1	M6 - FFG - T1	M5 - FFM - T2	M8 - FFA - T2
7	M4 - FFM - T1	M3 - FFG - T1	M2 - FFA - T2	M1 - FFA - T2	F8 - FFA - T2	F7 - FFG - T2	3 (F - P) - T1	F5 - FFM - T1
8	F4 - FFG - T2	F3 - FFG - T2	F2 - FFG - T1	F1 - FFA - T1	M8 - FFA - T1	M7 - FFM - T1	M6 - FFG - T2	M5 - FFM - T2

En cada casilla de grupo léase: identificación del locutor: sexo (masculino=M/femenino=F) + número del locutor dentro de su sexo – frecuencia fundamental (FFA=aguda/FFM=media/FFG=grave) – texto (texto 1= T1/texto 2=T2).

P=posición en el grupo de percepción.

ridad, honestidad, inteligencia, naturalidad, tranquilidad, profundidad, responsabilidad, seguridad, simpatía, sinceridad y verosimilitud. Todos estos indicadores fueron medidos con escalas bipolares de siete grados, según la técnica del “diferencial semántico”.

ANÁLISIS DE LOS DATOS

Procedimiento

El primer paso en el análisis fue explorar el comportamiento de la escala de las trece variables dependientes. Así, se obtuvo, después de diversas pruebas relacionadas con su estructura y mediante un análisis factorial de máxima verosimilitud con extracción oblicua, que el conjunto presentaba un carácter unidimensional (credibilidad) formado por dos subcomponentes altamente correlacionados entre ellos ($r=.840$). El test KMO de esfericidad resultó significativo (valor= .920; $p<.001$) y el goodness-of-fit test fue también significativo ($\chi^2=1682,96$; $p<.001$). La estructura extraída, una definición de credibilidad, se debía aceptar como altamente adecuada y confiable.

Dado este resultado, inmediatamente se procedió a revisar la literatura sobre medición de credibilidad (vgr. Berlo, Lemert y Mertz, 1969; Markham, 1968; McCroskey, 1966; McCroskey y Jenson, 1975; McCroskey y Young, 1981 y Whitehead, 1968). De estas obras, nos basamos en McCroskey y Young (1981:24-34), quienes reportan los hallazgos de más de una década de investigación experimental sobre la medición del constructo (sus propiedades y dimensiones) en distintas situaciones comunicativas. Basados en las conclusiones de estos investigadores, ampliamente fundamentadas, esta investigación decidió definir las dimensiones personalidad del locutor (PER) y competencia del locutor (COMP) para los dos factores arrojados por el análisis factorial. Según los resultados estadísticos, cada una de esas dos dimensiones se define por la acción de seis indicadores. Así, la primera dimensión, PER, hace referencia a indicadores relacionados con la habilidad y la experiencia para desarrollar un comportamiento o llevar a cabo una tarea; la segunda, COMP, se refiere a aspectos del carácter de la fuente y su adecuación a la tarea que desarrolla.

Por lo que respecta a los datos numéricos, la dimensión PER incluye a la naturalidad, verosimilitud, agrado, sinceridad, honradez y sim-

patía percibida en los hablantes (Alfa de Cronbach=.88) y explica un 51.44% de la varianza total. Por su parte, la dimensión COMP aglutina la seguridad, profundidad, tranquilidad, familiaridad, responsabilidad e inteligencia que los sujetos experimentales atribuyen a los locutores (Alfa de Cronbach=.82) y explica el 8.95% total de la varianza. Dada la fiabilidad de este modelo (Alfa de Cronbach total=.91), se decidió que ambas dimensiones serían exploradas en profundidad durante el estudio. Nótese, sin embargo, la diferencia entre el peso de ambos factores en el modelo y la superioridad de la dimensión PER como predictor del comportamiento del constructo.

A continuación, y una vez que fue definida la estructura factorial con la puntuación en credibilidad total (CT) y en cada uno de sus componentes (PER y COMP), se realizó un contraste de medias (ANOVA y MANOVA). Se compararon, entonces, las puntuaciones de las escalas aditivas en: 1) CT (la escala total); 2) PER y COMP (los sub-componentes o dimensiones de la Credibilidad); 3) CM (el promedio de enjuiciamiento del par creíble/No creíble de los cuestionarios) que se observó, únicamente, como dato orientador del promedio de la atribución de Credibilidad de los participantes, y 4) El resto de variables dependientes (que también, a partir de ahora, llamaremos *indicadores*) observados. Los grupos a los que se aplicó el contraste de medias fueron los definidos por las variables independientes (VI) y sus niveles: 1) frecuencia fundamental de la voz – FF- (aguda –FFA-, media –FFM- y grave –FFG-) y 2) modalidad de percepción (percepción sonora –PS- y percepción audiovisual –PA-).

Finalmente, tómesese en cuenta que se exploraron 2.560 evaluaciones distribuidas de la siguiente manera: 1) 640 evaluaciones correspondientes a los juicios de 80 sujetos sobre voces con FFG; 2) 1.440 evaluaciones correspondientes a los juicios de 180 sujetos sobre voces con FFM, y 3) 480 evaluaciones correspondientes a los juicios de 60 sujetos sobre voces con FFA.

RESULTADOS

A continuación se reporta el efecto de la FF sobre los indicadores, independientemente de si son únicamente escuchados los locutores (PS) o, además, se ven sus imágenes (PA). Una vez completado el análisis de

este factor, se reporta el análisis que toma en cuenta la interacción con la MP. Con el fin de facilitar la lectura y comprensión, los datos numéricos de este tratamiento se incluyen pormenorizados en la tabla 1.

TABLA 1
MEDIAS DE LAS VARIABLES DEPENDIENTES EN FUNCIÓN DE
LA FRECUENCIA FUNDAMENTAL DE LAS VOCES DE LOS LOCUTORES

	Frecuencia fundamental			<i>F</i>	<i>p</i>
	FFG	FFM	FFA		
	(<i>n</i> =640) M (<i>SD</i>)	(<i>n</i> =1440) M (<i>SD</i>)	(<i>n</i> =480) M (<i>SD</i>)		
Credibilidad total (CT)	47.92 (14.98)	58.23 (14.17)	52.74 (13.02)	122.30	< .001
Personalidad (PER)	23.57 (7.84)	28.64 (7.95)	27.08 (7.52)	92.43	< .001
Competencia (COMP)	24.36 (8.32)	29.60 (7.25)	25.70 (6.61)	128.69	< .001
Credibilidad media (CM)	4.04 (1.92)	5.05 (1.70)	4.70 (1.62)	74.50	< .001
Agrado	3.79 (1.82)	5.10 (1.67)	4.69 (1.69)	130.94	< .001
Familiaridad	3.48 (1.75)	4.57 (1.81)	4.21 (1.67)	81.69	< .001
Honradez	4.81 (1.45)	5.21 (1.34)	4.83 (1.33)	26.51	< .001
Inteligencia	3.83 (1.77)	4.59 (1.69)	4.45 (1.54)	46.16	< .001
Naturalidad	3.38 (1.95)	4.37 (1.98)	4.30 (1.88)	58.99	< .001
Profundidad	4.13 (1.98)	4.47 (1.75)	3.61 (1.57)	42.47	< .001
Responsabilidad	4.59 (1.60)	5.26 (1.34)	4.78 (1.36)	57.75	< .001
Seguridad	3.84 (2.10)	5.32 (1.73)	4.90 (1.71)	145.63	< .001
Simpatía	3.63 (1.54)	4.67 (1.62)	4.49 (1.62)	93.99	< .001
Sinceridad	4.12 (1.63)	4.69 (1.60)	4.38 (1.55)	26.77	< .001
Tranquilidad	4.49 (2.04)	5.39 (1.68)	3.74 (1.86)	167.33	< .001
Verosimilitud	3.86 (1.76)	4.60 (1.70)	4.40 (1.57)	42.77	< .001

FFG=frecuencia fundamental grave, FFM=frecuencia fundamental media y FFA=frecuencia fundamental aguda)

Los análisis estadísticos muestran que existen diferencias estadísticamente significativas en los promedios de todos los estimadores de credibilidad (CT, PER, COMP y CM) según la FF. Las voces con FFM son consideradas más creíbles que las voces con FFA. A su vez, estas últimas son consideradas más creíbles que las voces con FFG.

Asimismo, este modelo de enjuiciamiento (por el que las voces con FFM resultan más creíbles que las voces con FFA y estas estadísticamente más creíbles que las voces con FFG), también se presenta en la evaluación de los indicadores de seguridad, profundidad, agrado, sinceridad, tranquilidad y familiaridad. Así, las voces con FFM son consideradas significativamente más seguras, profundas, agradables, sinceras, tranquilas y familiares que las voces con FFA y estas, por su parte, estadísticamente, son consideradas más poseedoras de estas cualidades que las voces con FFG. Se determina, entonces, que estos indicadores son los que contribuyen a determinar las diferencias generales en la credibilidad de los locutores según sean las frecuencias de sus voces.

Además, el análisis de los datos detectó otro modelo de relación entre la credibilidad que obtienen las voces con FFM y FFA con respecto a las voces con FFG. Las FFM y FFA reciben superiores promedios de naturalidad, verosimilitud, inteligencia y simpatía que las voces con FFG. Luego, las diferencias entre la inferior credibilidad que las FG obtienen respecto a las FFM y FFA se debe a que los sujetos consideran que, estadísticamente, las voces con FFG son menos naturales, verosímiles, inteligentes y simpáticas que las FFM y FFA.

Por otra parte, en el análisis de los datos también se observó un tercer modelo presente en la relación que establecen los receptores respecto a las voces con FFM en relación con las que tienen FFA y las de FFG. Así, los datos muestran que las voces con FFM obtienen superiores promedios de responsabilidad y honradez que las otras frecuencias. Luego, la diferencia esencial entre la alta credibilidad que reciben las voces con FFM respecto al resto de las frecuencias, es que estas son vinculadas por los sujetos con la percepción de responsabilidad y honradez.

Una vez explorado el efecto de la FF sobre las variables dependientes e indicadores, se da cuenta, ahora, de los resultados de su interacción con la modalidad de percepción (MP). En este sentido, la

aplicación del modelo multivariado arrojó que existe interacción significativa entre MP y la FF en alguno de los estimadores e indicadores de la credibilidad y que, además, ambos factores, por separado, tienen efecto en varias de las medidas de la credibilidad (ver tablas 2 y 3 que incluyen los datos numéricos).

TABLA 2
MEDIAS DE LOS ESTIMADORES DE CREDIBILIDAD,
EN FUNCIÓN DE LA MODALIDAD DE PERCEPCIÓN Y LA
FRECUENCIA FUNDAMENTAL DE LAS VOCES

Prueba	CT	PER	COMP	CM
F	3.566	3.086	3.934	.793
p	0.28	.046	.020	.453
Modalidad sonora				
	M			
	(SD)			
FFG	49.166*** (.793)	24.006 (.439)	23.537 ** (.414)	4.085 (.098)
FFM	57.725*** (.529)	28.161 (.293)	29.615 (.276)	5.010 (.065)
FFA	52.351*** (.916)	27.013 (.507)	25.996 ** (.478)	4.619 (.113)
Modalidad audiovisual				
	M			
	(SD)			
FFG	46.672*** (.792)	23.134 (.438)	25.160** (.415)	4.762 (.112)
FFM	58.739*** (.528)	29.124 (.292)	29.564 (.277)	5.101 (.065)
FFA	53.133*** (.914)	27.137 (.506)	25.339** (.479)	3.994 (.097)

CT = credibilidad total; PER= personalidad; COMP= competencia; CM= credibilidad media; FFG=frecuencia fundamental grave; FFM=frecuencia fundamental media y FFA=frecuencia fundamental aguda ($p < .05$; ** $p < .01$; *** $p < .001$)

TABLA 3
VARIABLES DEPENDIENTES EN FUNCIÓN DE LA MODALIDAD DE PERCEPCIÓN Y LA FRECUENCIA FUNDAMENTAL DE LAS VOCES

Prueba	A	F	I	H	N	P	R	SE	SI	SIN	T	V
F	.547	.527	1.362	.116	5.411	4.120	2.108	3.267	4.346	1.241	18.093	3.149
p	.579	.591	.256	.890	.005	.016	.122	.038	.013	.289	.000	.043
Modalidad sonora												
M												
(SD)												
FFG	3.809 (.096)	3.470 (.099)	3.962 (.094)	4.856 (.076)	3.539 (.109)	4.295 (.099)	4.584 (.079)	4.019 (.102)	3.762 (.090)	4.144 (.090)	4.824 (.101)	3.997 (.094)
FFM	5.043 (.064)	4.584 (.066)	4.610 (.063)	5.292 (.051)	4.256 (.073)	4.401 (.066)	5.240 (.053)	5.291 (.068)	4.581 (.060)	4.598 (.060)	5.395 (.067)	4.546 (.063)
FFA	4.632 (.111)	4.297 (.114)	4.569 (.109)	4.912 (.088)	4.418 (.126)	3.548 (.115)	4.908 (.091)	4.866 (.118)	4.439 (.103)	4.335 (.103)	3.410 (.116)	4.448 (.109)
Modalidad audiovisual												
M												
(SD)												
FFG	3.762 (.096)	3.506 (.099)	3.681 (.094)	4.755 (.077)	3.219 (.109)	3.956 (.099)	4.589 (.079)	3.647 (.102)	3.494 (.089)	4.094 (.089)	4.163 (.101)	3.997 (.094)
FFM	5.158 (.064)	4.551 (.066)	4.567 (.063)	5.137 (.051)	4.478 (.073)	4.529 (.066)	5.284 (.053)	5.346 (.068)	4.758 (.060)	4.783 (.060)	5.382 (.067)	4.546 (.063)
FFA	4.754 (.110)	4.117 (.114)	4.329 (.109)	4.741 (.089)	4.175 (.126)	3.671 (.115)	4.649 (.091)	4.921 (.118)	4.537 (.103)	4.404 (.103)	4.050 (.116)	4.448 (.109)

A=agrado; F=familiaridad; I=inteligencia; H=honestidad; N=naturalidad; P=profundidad; R=responsabilidad;

SE=seguridad; SI=simpatía; SIN=sinceridad; T=tranquilidad; V=verosimilitud; FFG=frecuencia fundamental grave; FFM=frecuencia fundamental media; FFA=frecuencia fundamental aguda.

Por una parte, se ha encontrado que las interacciones significativas entre estas dos variables (MP y FF) se encuentran en CT, las dimensiones PER y COMP (tabla 2) y en los indicadores de seguridad, naturalidad, profundidad, verosimilitud, tranquilidad y simpatía (tabla 3). Se explicarán uno a uno los detalles de estas interacciones.

El análisis de la actuación en conjunto de la MP y la FF sobre la CT muestra que cuando los sujetos únicamente escuchan las voces de los hablantes (PS) tienden a atribuir promedios más semejantes entre las propiedades de la FF que cuando también ven sus imágenes. Así, los promedios de CT que obtienen las voces con FFG, FFM y FFA son mucho más diferentes (se alejan) en PA que en PS. Luego, la MP afecta a la definición de las diferencias entre los promedios de credibilidad de las voces de todas las FF. Mediante las imágenes de los locutores se amplían, reconocen, o expanden las impresiones que sobre la credibilidad de esas frecuencias obtienen las voces (tabla 2).

Con respecto a la actuación de la MP y la FF en conjunto sobre la dimensión PER, el análisis muestra que los juicios a las variables que forman parte de esta dimensión tienden a diferenciarse menos en PS y a diferenciarse más en PA. Luego, de la misma manera que sucede con el estimador CT, la percepción de las imágenes de los locutores ocasiona que los juicios sobre la dimensión PER de la credibilidad se amplíen, magnifiquen o distingan en mayor medida que cuando únicamente se escuchan sus voces. Por otra parte, valga decir que en ambas MP, la voz con FFM obtiene los superiores promedios de la dimensión PER y la voz con FFG obtiene los inferiores promedios en dicha dimensión (tabla 2).

Con respecto a la actuación de la MP y la FF en conjunto sobre la otra dimensión de la credibilidad, la competencia, COMP, el análisis muestra que la voz con FFM recibe los superiores promedios de esta dimensión en ambas MP. Tanto cuando se escucha a las voces como cuando se ven sus imágenes, los locutores con FFM son considerados más competentes. La relación entre los promedios de COMP de FFG y FFA es semejante en PS, pero se distingue o distancia en PA debido a que las voces con FFA reciben superiores promedios de COMP (tabla 2).

A partir de los datos de los párrafos anteriores se puede concluir que la PA afecta, en sentido negativo, y en superior proporción, a las evalua-

ciones de PER y COMP de las voces con FFG. Estas, en consecuencia, son las que sufren los efectos más claros de la MP.

A continuación, tomamos en cuenta al indicador tranquilidad (que tiene la interacción más significativa del análisis), que es parte de la dimensión COMP. Así, se observa que en ambas MP, las voces con FFM son las que obtienen los superiores promedios de ese indicador (tabla 3). El análisis muestra, además, que la relación entre los promedios de tranquilidad entre las voces FFG y las FFA (es decir, las frecuencias extremas) es más evidente, separada y disímil en PS que en PA. Con la percepción de las imágenes de los locutores, luego, la distancia entre los promedios de tranquilidad que reciben las voces FFA y FFG se reduce.

Por otra parte, se observa que la interacción de la FF y la MP afecta al indicador naturalidad, en segundo lugar de efecto (tabla 3). Asimismo, se observa que esta es el indicador de la dimensión PER que recibe los mayores resultados de la acción de ambas variables. El análisis muestra que los promedios de naturalidad atribuidos a las voces con FFA y FFG se aproximan en PS y se disparan, se diferencian o se distinguen en PA.

El indicador simpatía también recibe los efectos de la interacción entre la FF y la MP (tabla 3). Los promedios de todas las FF se aproximan en PS (en particular los de las FFM y FFA que reciben superior valoración que los de las FFG) y se separan distinguen o amplían en PA siguiendo la misma tendencia (FFM y FFA con superiores promedios y FFG con muy inferiores).

Asimismo, la interacción afecta al indicador profundidad en dos sentidos (tabla 3). Por una parte, la profundidad de las voces FFM y FFG se aproximan en PS y se disparan o distinguen en PA y en especial con las voces FFM. También, en un segundo sentido, las evaluaciones sobre Profundidad de las voces con FFA y FFG se aproximan en PA y se distinguen en PS con superiores promedios para las voces FFG.

Por lo que respecta al indicador seguridad, el análisis muestra que las propiedades de las frecuencias media fundamentales se aproximan en PS y se distancian en PA. Esa diferenciación supone que las voces con FFG reciban menores promedios de seguridad y las FFM y FFA superiores (tabla 3).

Finalmente, la interacción afecta al indicador verosimilitud. La interacción sigue la siguiente dirección: las tres voces se aproximan en PS

(en particular FFM y FFA) y se distancian en PS (hasta que queda muy por debajo la verosimilitud de la voz con FFG en PS) (tabla 3).

En conclusión, el análisis muestra que la MP, por sí sola, únicamente afecta significativamente a la inteligencia y la honradez, mientras que la FF, por sí sola, afecta a todas las variables dependientes. Sin embargo, y en resumen, la interacción (modalidad x frecuencia) es significativa en las diferencias de media para: credibilidad total, dimensiones de personalidad y competencia y los siguientes indicadores en este orden de significación (de mayor a menor influencia): tranquilidad, naturalidad, simpatía, profundidad, seguridad y verosimilitud.

CONCLUSIONES

En el punto anterior dimos cuenta completa de los resultados del tratamiento experimental. A continuación, observamos sus aspectos más relevantes.

Los resultados reportados confirman las dos hipótesis de este trabajo. Por una parte, prueban que la frecuencia fundamental de las voces es un elemento influyente en la formación de impresiones de las audiencias sobre los intérpretes mediáticos. Por otra, muestran que existe un efecto de la interacción conjunta de la frecuencia fundamental y la percepción audiovisual en la formación de impresiones. De hecho, la percepción del rostro afecta, en distinta medida, a los juicios de los receptores dependiendo de si son hablantes con voces agudas, medias y graves.

Estos resultados añaden vigor a la hipótesis de la dominancia de la señal sonora del habla (la voz) sobre la visual (el rostro) en la formación de impresiones sobre los intérpretes audiovisuales (Daly y Bench, 1996; Massaro y Cohen, 1996), añaden datos al gran cuerpo de evidencias sobre la integración perceptiva sonoro-visual (MacDonald y McGurk, 1978; Massaro y Cohen, 1983, 1996; Summerfield, 1987), y son otro alegato a favor del estudio de la voz en comunicación mediática. Por ello, creemos oportuno que futuros estudios profundicen en las cualidades del habla apropiada para los medios, en particular de las señales no verbales (*vgr.* velocidad, ritmo, intensidad, manipulación del timbre, articulación, acento) y su relación con la formación de impresiones de las audiencias en distintos contextos comunicativos.

No obstante la dominancia sonora sobre la visual, y porque los resultados han mostrado que el efecto de la modalidad de percepción es superior con las voces de frecuencias extremas (graves o agudas), parecería que la percepción del rostro (quizás por la presencia de los movimientos articulatorios del habla, en atención a lo que dicen expertos en percepción del lenguaje) ejerce una función clarificadora de la formación de impresiones sobre esas voces. En consecuencia, una de las funciones de la percepción audiovisual del rostro sería mejorar la evaluación del habla aportando inteligibilidad al enjuiciamiento de las voces con frecuencias distanciadas de la media.

Sin embargo, merece destacar que este estudio también ha encontrado que existe mayor variabilidad en los juicios de los receptores cuando se *audiove* a los locutores (como en televisión) que cuando se les escucha (como en radio), por lo que el enjuiciamiento sonoro es más consistente entre las audiencias que el enjuiciamiento audiovisual. Sin embargo, este dato debe ser tomado con precaución: la mayor variabilidad del enjuiciamiento audiovisual frente al sonoro se refleja en la existencia de mayores diferencias entre las distintas propiedades evaluadas. Luego, la percepción audiovisual del rostro de los locutores contribuiría a que se magnificaran, evidenciaran o clarificaran los contrastes de algunas de las impresiones sobre los distintos tipos de voces. Así, en televisión, las diferencias en los promedios de enjuiciamiento a los distintos tonos de voces de los hablantes tenderían a ser siempre superiores a las diferencias de enjuiciamiento a los distintos tonos en radio.

Por otra parte, y de acuerdo a los datos aportados por Strand (1999:86-101) y Wyer *et al.* (1994:254-267), las mayores diferencias en los enjuiciamientos en las pruebas audiovisuales, con respecto a las sonoras, se podrían explicar por el hecho de que la información visual de los hablantes envuelve el acceso a expectativas sociales sobre los hablantes en función de cómo lucen. Futuros estudios, no obstante, deberían profundizar, específicamente, en los elementos de la imagen que contribuyen a obtener superior variabilidad de los juicios sobre las voces de los hablantes mediáticos.

Asimismo, los resultados aportados parecen apoyar a Treviño, Lengel y Draft (1987:553-575) quienes, a la luz de la *richness media theory*, consideran que las impresiones de las audiencias de los medios

de comunicación más ricos en información son más precisas; que la información visual tiende a homogeneizar las percepciones. Este estudio muestra que una proporción de las percepciones promedios sobre los hablantes podrían ser más definitivas (más claras) cuando se dispone de la imagen (de las claves visuales). No obstante, los resultados confirman que las audiencias extraen las impresiones sobre el hablante fundamentalmente a partir de los elementos sonoros de su habla y que el enjuiciamiento promedio es más homogéneo en el medio “menos rico” (en el sonoro). Esto conduce a pensar, por una parte, que la homogeneidad de las impresiones es un factor distinto al de precisión del enjuiciamiento. Es decir, que una cosa es que las audiencias estén más o menos de acuerdo sobre sus impresiones acerca de un actor mediático y otra que sus juicios sean más claros, definitivos o extremos. Esos datos, además, también conducen a afirmar que las audiencias se forman impresiones sobre los actores mediáticos a partir de la obtención de la información mínima necesaria para dar cuenta de la tarea que se le ha encomendado hacer, es decir, de forma pragmática y orientada, como dirían Abele y Petzold (1998). Finalmente, llevan a hipotetizar que la eficacia de la comunicación de las organizaciones que emiten mensajes sonoros y audiovisuales (vgr. las grandes corporaciones y a la luz de las evidencias acerca de un crecimiento de productos *cross-media*) podría depender de la mejor explotación de los elementos pertinentes al medio menos rico en la configuración del mensaje que envían a sus audiencias. Estudios posteriores deberían probar estas suposiciones.

Por otra parte, este trabajo también encuentra que la credibilidad es el factor que guía, esencialmente, la formación de impresiones sobre los actores de la representación mediática y es el elemento que ayuda a organizar el resto de las evaluaciones de las audiencias. Estos resultados son coherentes con los reportados por los investigadores más prominentes de la credibilidad mediática (Berlo, Lemert y Mertz, 1969; Markham, 1968; McCroskey, 1966; McCroskey y Jenson, 1975; McCroskey y Young, 1981 y Whitehead, 1968), quienes defienden que es la variable más influyente en el enjuiciamiento de las personalidades que aparecen en los medios. Pero, asimismo, los datos aportados por este estudio señalan que la atribución de credibilidad se desprende de las percepciones sobre la PER y la competencia percibida (con mayor peso de la primera que de

la segunda), lo que confirma los hallazgos del programa experimental de James C. McCroskey y sus asociados, que demuestran que en la base de la definición de credibilidad se encuentran ambas dimensiones.

Los resultados de este estudio son aplicables al diseño de contenidos audiovisuales más eficaces porque proveen información útil (y basada en el estudio de sus efectos sobre las audiencias) para el control y la manipulación expresiva de la voz o de la imagen. Creemos, asimismo, que los resultados podrían ser sugerentes para los interesados en profundizar sobre los procesos perceptivos o de la formación de impresiones de las audiencias de radio y televisión.

Bibliografía

- ABELE, Andrea E., Petzold, Peter (1998) "Pragmatic use of categorical information in impression formation", *Journal of Personality and Social Psychology*, vol. LXXV, núm. 2.
- AMBADY, Nailini, Bernieri, Frank J., Richeson, Jennifer A. (2000) "Toward a histology of social behaviour: judgemental accuracy from think slices of the behavioural stream", *Advances in Experimental Social Psychology*, vol. XXXII, .
- APPLE, William, Streeter, Lynn A., Krauss, Robert M. (1979) "Effects of pitch and speech rate on personal attributions", *Journal of Personality and Social Psychology*, vol. XXXVII, núm. 5.
- BARGE, J. Kevin, Schlueter, David W., Pritchard, Alex (1989) "The effects of nonverbal communication and gender on impression formation in opening statements", *Southern Communication Journal*, vol. LIV, summer.
- BERLO, David K., Lemert, James B., Mertz, Robert J. (1970) "Dimensions for evaluating the acceptability of message sources", *Public Opinion Quarterly*, vol. XXXIII, núm. 4.
- BOND, Ronald N., Welkowitz, Joan., Goldschmid, Harlene, Wattenberg, Sarah (1987) "Vocal frequency and person perception: effects of perceptual salience and nonverbal sensitivity", *Journal of Psycholinguistic Research*, vol. XVI, núm. 4.
- BURGOON, Judee K., Birk, Thomas, Pfau, Michael, (1990) "Nonverbal behaviors, persuasion, and credibility", *Human Communication Research*, vol. XVII, núm. 1.

- (1978) "Attributes of the newscaster's voice as predictors of his credibility, *Journalism Quarterly*, summer.
- CARUSO, Anthony, Mueller, Peter B., Shadden, Barbara B., (1995) "Effects of aging on speech and voice", *Pathology and Audiology*, vol. XIII, núm. 2.
- COLLINS, Sarah A., (2000) "Men's voices and women's choices", *Animal Behaviour*, vol. LX, núm. 6.
- COLLINS TIDWELL, Lisa, Walther, Joseph B., (2002) "Computer-mediated communication effects on disclosure, impressions, and interpersonal evaluations: getting to know one another a bit a time", *Human Communication Research*, vol XXVIII, núm. 3.
- DAFT, Richard L., Lengel, Robert H., (1984) "Information richness: A new approach to managerial behaviour and organizational design", *Research in Organizational Behaviour*, vol. VI.
- DALY, Nicola., Bench, John, (1996) "Interpersonal impressions, gender stereotypes and visual speech", *Journal of Language and Social Psychology*, vol. XV, núm. 4.
- De PAULO, Bella M., Friedman, Howard S. (1998) "Nonverbal communication", en Daniel T. Gilbert, Susan T. Fiske y Gardner Lindzey (eds.), *The handbook of social psychology* (4th ed., vol. II), New York: McGraw Hill.
- DUNNING, David, Griffin, Dale W., ROSS, Lee, Milojkovic, James D., (1990) "The overconfidence effect in social prediction", *Journal of Personality and Social Psychology*, vol. LVIII, núm. 4.
- ERICKSON, Bonnie, Lind, E. Allan, Johnson, Bruce C., (1978) "Speech style and impression formation in a court setting: the effects of 'powerful' and 'powerless' speech", *Journal of Experimental Social Psychology*, vol. XIV, núm. 3.
- FAIRBANKS, Grant, (1940) "Recent experimental investigations of vocal pitch in speech", *Journal of the Acoustical Society of America*, vol. XI.
- Pronovost, Wilbert (1939) "An experimental study of the pitch characteristics of the voice during the expression of emotion", *Speech Monographs*, vol. VI, núm. 1.
- FISKE, Susan T., Neuberg, Steven L. (1990) "A continuum of impression formation, from category-based to individuating processes:

- Influences of information and motivation on attention and interpretation". *Advances in Experimental Social Psychology*, vol. XXIII.
- FISKE, Susan T., Taylor, Shelley.E. (1991) *Social cognition*, New York: McGraw-Hill.
- HIGGINS, E. Tory, Rholes, William S., Jones, Carl R. (1977) "Category accessibility and impression formation", *Journal of Experimental Social Psychology*, vol. XIII, núm. 2.
- JACOBSON, David, (1999) "Impression formation in cyberspace: online expectations and offline experiences in text-based virtual communities", *Journal of Computer Mediated Communication*, vol. V, núm. 1, <http://jcmc.indiana.edu/vol5/issue1/jacobson.html> (último acceso el 28 de agosto de 2007).
- LAY, Clarry H., Burron, Bryan F. (1968) "Perception of the personality of the hesitant speaker", *Perceptual and Motor Skills*, vol. XXVI, núm. 3.
- LAUKKA, Petri, Juslin, Patrik, Bresin, Roberto (2005) "A dimensional approach to vocal expression of emotion", *Cognition and Emotion*, vol. XIX, núm. 5.
- MACDONALD, John, MCGURK, Harry (1978) "Visual influences on speech perception processes", *Perception and Psychophysics*, vol. XXIV, núm. 3.
- MARKHAM, David, (1968) "The dimensions of source credibility of television newscasters", *Journal of Communication*, vol. XVIII, march.
- MASSARO, Dominic W., Cohen, Michael M. (1996) "Perceiving speech from inverted faces", *Perception and Psychophysics*, vol. LVIII, núm. 7.
- (1983) "Evaluation and integration of visual and auditory information in speech perception", *Journal of Experimental Psychology: human perception and performance*, vol. IX, núm. 5.
- MCCROSKEY, James C., (1966) "Scales for the measurement of ethos", *Speech Monographs*, vol. XXXIII, núm. g.
- Mehrley, R. Samuel (1969) "The effects of disorganization and nonfluency on attitude change and source credibility", *Speech Monographs*, vol. XXXVI, núm. 1.
- Jenson, Thomas A. (1975) "Image of mass media news sources", *Journal of Broadcasting*, vol XIX, núm. 19.

- Young, Thomas J. (1981) “Ethos and credibility: the construct and its measurement after three decades”, *The Central States Speech Journal*, vol. XXXII.
- MILLER, Gerard R., Hewgill, Murray A. (1964) “The effect of variations in nonfluency on audience ratings of source credibility”, *Quarterly Journal of Speech*, vol. L, núm. 1.
- MONTEPARE, Joann M., Zebrowitz-McCarthur, Leslie (1989) “Contributions of a babyface and a childlike voice to impressions of moving and talking faces”, *Journal of Nonverbal Behaviour*, vol. XIII, núm. 3.
- NIEDZIELSKI, Nancy (1999) “The effect of social information on the perception of sociolinguistic variables”, *Journal of Language and Social Psychology*, vol. XVIII, núm. 1.
- NEWMAN, Leonard S., Uleman, James (1990) “Assimilation and contrast effects in spontaneous trait inference”, *Personality and Social Psychology Bulletin*, vol. XVI, núm. 2.
- OSTERMEIER, Terry H. (1967) “Effects of type and frequency of reference upon perceived source credibility and attitude change”, *Speech Monographs*, vol. XXXIV, núm. 2.
- PRADO, Emili (investigador principal) (1992) “Gramática de la expresión fonostésica y la representación imaginativo-visual en los sistemas integrales de percepción de la voz” (informe final de la investigación PB88-0236 de la DGICYT), Departamento de Comunicación Audiovisual y Publicidad, Universidad Autónoma de Barcelona.
- PRADO, Emili (investigador principal) (1997) “El modelo acústico MAC de la voz en el marco de la gramática de la expresión fonostésica” (informe final de la investigación PB93-0850 de la DGICYT), Departamento de Comunicación Audiovisual y Publicidad, Universidad Autónoma de Barcelona.
- RIDING, David, Lonsdale, Deryle, Brown Bruce, “The effect of average fundamental frequency and variance of fundamental frequency on male vocal attractiveness to women”, *Journal of Nonverbal Behaviour*, vol. XXX, núm. 2.
- RODRÍGUEZ, B., Angel (1989) “La construcción de una voz radiofónica”, tesis doctoral, Universidad Autónoma de Barcelona, Departamento de Comunicación Audiovisual y Publicidad, Bellaterra.

- SCHERER, Klaus R. (1979) "Personality markers in speech", en Scherer, K.R., Giles, H. (eds.), *Social markers in speech*, New York: Cambridge University press.
- (1978) "Personality inference from voice quality: the loud voice of extroversion", *European Journal of Social Psychology*, vol. VIII, núm.4.
- London, Harvey, Wolf, Pared J. (1973) "The voice of confidence: Paralinguistic cues and audience evaluation", *Journal of Research in Personality*, vol. VII, núm. 1.
- Scherer, Ursula, Hall, Judith, A. (1977) "Differential attribution of personality based on multi-channel presentation of verbal and non-verbal cues", *Psychological Research*, vol. XXIX, núm. 3.
- SHERMAN, Richard C., End, Christian, Kraan, Egon, Cole, Alison, Campbell, Jammon, Klausner, Jaime, Birchmeier, Zachary (2001) "Metaperception in cyberspace", *Cyberpsychology Behaviour*, vol IV, núm. 1.
- SERENO, Kenneth K., Hawkings, Gary, J. (1967) "The effects of variations in speakers nonfluency upon audience ratings of attitude toward the speech topic and speaker's credibility", *Speech Monographs*, vol. XXXIV, núm. 1.
- SIEGMAN, Aron W., Reynolds, Mark. A. (1983) "Effects of mutual invisibility and topical intimacy on verbal fluency in dyadic communication", *Journal of Psycholinguistic Research*, vol. XII, núm. 5.
- STRAND, Elizabeth A. (1999) "Uncovering the role of gender stereotypes in speech perception", *Journal of Language and Social Psychology*, vol. XVIII, núm. 1.
- STREET, Richard L., (1982) "Evaluation of noncontent speech accommodation", *Language & Communication*, vol. XXI, núm. 1.
- SOTO, María Teresa (2000) "Influencia de la percepción visual del rostro del hablante en la credibilidad de su voz", tesis doctoral, Departamento de Comunicación Audiovisual y Publicidad, Universidad Autónoma de Barcelona, España.
- SUMMERFIELD, Quentin (1987) "Some preliminaries to a comprehensive account of audio-visual speech perception", en Barbara Dodd y Ruth Campbell (eds.), *Hearing by eye: the psychology of lip-reading*, Hillsdale: Lawrence Erlbaum Associates.

- TREVIÑO, Linda K., Lengel, Robert H., Daft, Richard L. (1987) "Media symbolism, media richness, and media choice in organizations", *Communication Research*, vol. XIV, núm. 5.
- ULEMAN, James, Newman, Leonard S. (1992) "Can personality traits be inferred automatically? Spontaneous inferences require cognitive capacity at encoding", *Consciousness and Cognition: an International Journal*, vol. I, núm. 1.
- WALTHER, Joseph B. (1993) "Impression development in computer-mediated interaction", *Western Journal of Communication*, vol. LVII.
- (1996) "Computer-mediated communication: impersonal, interpersonal, and hyperpersonal interaction", *Communication Research*, vol. XXIII, núm. 1.
- WEISBAND, Suzanne P., Weisband, Sherry K., Schneider, Terry C., (1995) "Computer-mediated communication and social information: status salience and status differences", *The Academy of Management Journal*, vol. XXXVIII, núm. 4.
- WHITEHEAD, Jr., Jack L. (1968) "Factors of source credibility", *Quarterly Journal of Speech*, vol. LIV, núm. 1.
- WYER, Robert S., Budesheim, Thomas Lee, Lambert, Alan J., Swan, Suzanne (1994) "Person memory and judgment: pragmatic influences on impressions formed in a social context", *Journal of Personality and Social Psychology*, vol. LXVI, núm. 2.
- ZUCKERMAN, Miron, Miyake, Kunitake (1993) "The attractive voice: what makes it so?", *Journal of Nonverbal Behaviour*, vol. XVII, núm. 2.