



## La cara auditiva: El reconocimiento de las personas a través de la voz

Alba Mónica Casado Ramos

Programa de Doctorado en Psicología, Universidad de Granada, España

Tipo de artículo: Actualidad.

Disciplinas: Psicología, Neurociencias.

Etiquetas: identificación del hablante, lenguaje, voz, cara auditiva, integración multisensorial.

*Cuando escuchamos a alguien hablar, aunque no lo veamos ni conozcamos, obtenemos diferentes tipos de información a partir de su voz: su mensaje lingüístico, su estado emocional y también, por extraño que parezca, una imagen de sus características físicas, una especie de "cara auditiva". Si el hablante es alguien conocido, podemos además saber quién es, es decir, podemos identificarlo únicamente por la idiosincrasia de su voz. Cuando voz y cara aparecen juntas, la información obtenida a partir de ambas modalidades sensoriales se integra para crear una imagen global y completa del hablante.*



(cc) MarkGregory007

La voz humana es el sonido más importante de nuestro entorno, y posiblemente el que más escuchemos a lo largo de nuestra vida. A nivel cerebral hay un conjunto de neuronas que se activan selectivamente con sonidos vocales (independientemente de que contengan o no información semántica) situadas en lo que se denomina "áreas temporales de voces" (TVA, siglas en inglés), a lo largo de las partes anteriores y mediales del giro y surco temporal superior (GTS/STS) en ambos hemisferios (Belin, Bestelmeyer, Latinus y Watson, 2011). Las técnicas electrofisiológicas también aportan evidencia de la existencia de una actividad neuronal específica para la voz, conocida como "respuesta sensible a la voz", que se da ante sonidos vocales y no ante otros sonidos naturales (Charest y col., 2009, citado en Belin y col., 2011).

La voz, junto con información lingüística, transmite también información ligada a las características físicas y emocionales del hablante. Por ello, la voz no es sólo el principal y más primitivo vehículo del lenguaje, sino que además es nuestra "cara auditiva". Cada hablante tiene

unas características acústicas específicas que vienen marcadas por la forma de su aparato vocal. Cuando el aire pasa a través de las cuerdas vocales de la laringe, éstas vibran produciendo un sonido complejo con una frecuencia fundamental (F0) que depende de la longitud y el grosor de las mismas. Este sonido básico va resonando a través de las cavidades superiores -faringea, oral y nasal-, concentrando la energía sonora en determinadas bandas de frecuencias y dando como resultado los formantes (F1, F2, F3,...). Por lo general, la voz de las mujeres presenta un patrón de frecuencias o timbre más alto que el de los hombres, tanto en su F0 (entre 165 y 225 Hz para las mujeres, frente a entre 85 y 180 Hz para los hombres) como en sus correspondientes formantes. Por lo tanto, analizando estos parámetros podemos identificar una serie de características físicas del hablante (edad, género, altura, peso...) y generar una imagen visual a partir de la voz.

Si el hablante es una persona conocida la podemos además identificar a través de su voz, activando la imagen visual que de ella tenemos almacenada en la memoria. Esta habilidad está presente desde que somos pequeños, aproximadamente a los 7 meses de edad, y se da también en otras especies animales, como los macacos (Belin y col., 2011). La habilidad de reconocer a las personas a través de la voz puede verse afectada por un déficit neurológico denominado fonoagnosia, que consiste en la incapacidad de discriminar entre diferentes voces y la incapacidad de reconocer voces familiares. Este déficit es análogo a la

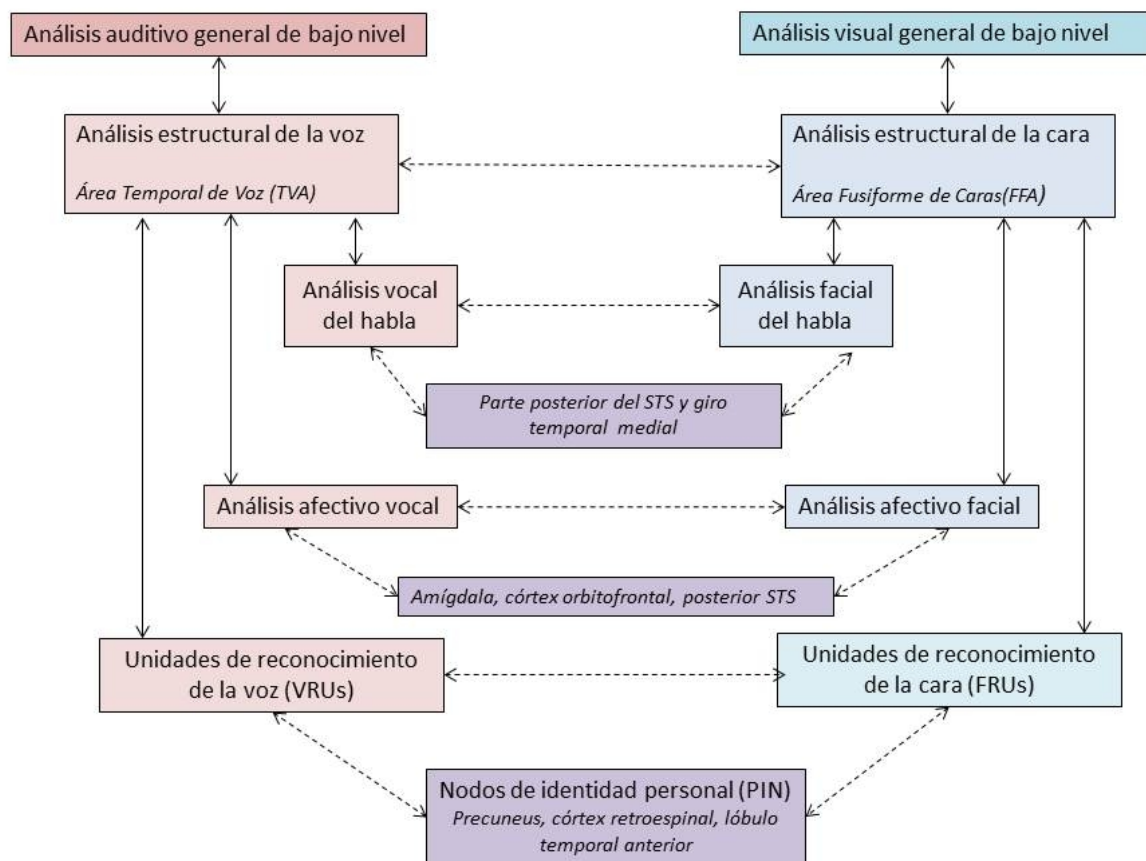


Figura 1.- Modelo de percepción de la voz de Belin, Fecteau y Bedard (2004). El procesamiento de la información vocal estaría dissociada en tres sistemas independientes: el que analiza la información semántica, el que analiza la información afectiva y el sistema que analiza la identidad del hablante. Estos sistemas entrarían en constante interacción con sus homólogos en el procesamiento de las caras. En la figura se mencionan también las principales áreas cerebrales implicadas en cada una de estas fases de procesamiento.

prosopagnosia, la incapacidad de reconocer a las personas a partir de su cara, y demuestra que el procesamiento de la identidad de las personas puede realizarse por diferentes vías independientes (Belin y col., 2011).

No obstante, en la mayoría de interacciones sociales comunicativas la información auditiva y visual se presentan juntas. De estas dos modalidades, los elementos clave para realizar la identificación del hablante son la voz y la cara. Supone una ventaja de procesamiento integrar estas dos fuentes para obviar la información redundante y maximizar la información recogida (Campanella y Belin, 2007). El modelo de reconocimiento de las personas a través de la voz propuesto por Belin, Fecteau y Bedard (2004) se inspira en el de Bruce y Young (1986) para el procesamiento de caras y da cuenta de la interacción entre la información auditiva y visual (véase la Figura 1).

Hay estudios conductuales que apoyan la interacción audiovisual. Por ejemplo, una cara familiar se reconoce antes cuando es precedida brevemente por su voz, frente a cuando la voz no es suya, y viceversa. También, cuando nos enfrentamos a la tarea de escuchar una voz y decidir a cuál de dos caras le corresponde, acertamos sin problema. De nuevo, esto también sucede a la inversa. Además, cuando vemos y escuchamos a alguien hablar, recordamos mejor quién es que si únicamente le escuchamos hablar (Campanella y Belin, 2007).

El debate actual se centra en conocer el momento de la integración de la información procedente de las diferentes modalidades: temprano, en los estados iniciales del procesamiento, o tardío, en las zonas de convergencia supramodales. Quiñones-González y col. (2011) realizaron un estudio de electroencefalografía (potenciales evocados) para aclarar la problemática y comprobaron que la integración se realiza a todos los niveles, desde los primeros estadios (unidades de reconocimiento de voces y caras) hasta los nodos de información de la persona, tal y como predijeron Belin y col. (2004) en su modelo.

En conclusión, con la información acústica propia de cada persona, que depende de su aparato fonatorio, creamos su “cara auditiva”, es decir, obtenemos información de cómo es físicamente, además de información acerca de su estado emocional. Cuando esa persona es conocida, escuchar su voz permite también acceder a la información almacenada sobre ella e identificarla, aunque no la veamos. Además, en los procesos de comunicación habituales integramos la información visual y auditiva disponible sobre el hablante, lo que redundará en una mayor eficacia y rapidez en su identificación.

## Referencias

- Belin, P., Fecteau, S., y Bedard, C. (2004). Thinking the voice: Neural correlates of voice perception. *Trends in Cognitive Sciences*, 8, 129–135.
- Belin, P., Bestelmeyer, P. E. G., Latinus, M., y Watson, R. (2011). Understanding voice perception. *British Journal of Psychology*, 102, 711-725.
- Bruce, V., y Young, A. (1986) Understanding face recognition. *British Journal of Psychology*, 77, 305–327.
- Campanella, S., y Belin, P. (2007). Integrating face and voice in person perception. *Trends in Cognitive Sciences*, 11, 535–543.
- Quiñones-González, I., Bobes-León, M. A., Belin, P., Martínez-Quintana Y., Galán-García, L., y Sánchez-Castillo, M. (2012). Person identification through faces and voices: An ERP study. *Brain Research*, 1407, 13-26.

Manuscrito recibido el 29 de enero de 2014.

Aceptado el 23 de febrero de 2014.