

PABLO QUINTANILLA | CARLA MANTILLA | PAOLA CÉPEDA  
(editores)

# COGNICIÓN SOCIAL Y LENGUAJE

La intersubjetividad en la evolución de la especie  
y en el desarrollo del niño



## Capítulo 17



FONDO  
EDITORIAL

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ

*Cognición social y lenguaje*

*La intersubjetividad en la evolución de la especie y en el desarrollo del niño*

Pablo Quintanilla, Carla Mantilla y Paola Céspedes (editores)

© Pablo Quintanilla, Carla Mantilla y Paola Céspedes, 2014

© Fondo Editorial de la Pontificia Universidad Católica del Perú, 2014

Av. Universitaria 1801, Lima 32, Perú

Teléfono: (51 1) 626-2650

Fax: (51 1) 626-2913

feditor@pucp.edu.pe

www.pucp.edu.pe/publicaciones

Diseño, diagramación, corrección de estilo  
y cuidado de la edición: Fondo Editorial PUCP

Primera edición: mayo de 2014

Tiraje: 500 ejemplares

Prohibida la reproducción de este libro por cualquier medio, total o parcialmente,  
sin permiso expreso de los editores.

Hecho el Depósito Legal en la Biblioteca Nacional del Perú N° 2014-07083

ISBN: 978-612-4146-80-0

Registro del Proyecto Editorial: 11501361400359

Impreso en Tarea Asociación Gráfica Educativa  
Pasaje María Auxiliadora 156, Lima 5, Perú

## EL CARÁCTER MENTAL DE LOS SISTEMAS FONOLÓGICOS

Jorge Iván Pérez Silva

Pontificia Universidad Católica del Perú

En su relato «Funes el memorioso», Jorge Luis Borges concibe un personaje dueño de una percepción y una memoria infalibles. He aquí algunos ejemplos de lo que Ireneo Funes era capaz de hacer: «Nosotros, de un vistazo» —dice Borges— «percibimos tres copas en una mesa; Funes, todos los vástagos y racimos y frutos que comprende una parra»<sup>1</sup>. La *simplicidad* de la imagen visual de tres copas contrasta notoriamente con la *complejidad*, con el abarrotamiento, de una parra. Para nosotros, con solo darle un vistazo, es casi imposible decir cuántas uvas hay inclusive en un racimo; Funes, en cambio, captaba nítidamente los innumerables individuos que componen toda una parra. Aquí otro ejemplo de la extraordinaria percepción de Funes: «Una circunferencia en un pizarrón, un triángulo rectángulo, un rombo, son formas que podemos intuir plenamente; lo mismo le pasaba a Ireneo con las aborascadas crines de un potro, con una punta de ganado en una cuchilla, con el fuego cambiante y con la innumerable ceniza». Comparemos una vez más nuestra percepción plena de una figura geométrica *simple* dibujada en un pizarrón, con la percepción que tenía Funes de una realidad *compleja*, abigarrada de detalles, como las crines de un caballo. Quizá las personas miopes podemos imaginar cómo percibía Ireneo Funes: al sacarnos los anteojos, nuestra imagen del mundo se torna borrosa; las crines de un caballo son indistinguibles y lo mismo ocurre con las hojas de los árboles; al ponernos los anteojos, nuestra imagen del mundo pierde su carácter de pintura impresionista —digamos— para adquirir algo de nitidez; yo imagino la visión de Funes como la que nos darían unos anteojos potentísimos, que no dejaran pasar desapercibido ningún detalle de los objetos, como nos ocurre a simple vista con una circunferencia o un triángulo en un pizarrón.

---

<sup>1</sup> Todas las citas de «Funes el memorioso» son de Borges (1983). No indico la página de cada cita para facilitar la lectura. El lector interesado podrá encontrar fácilmente en cualquier edición los fragmentos citados dada la brevedad del relato.

Lo que busca Borges con esta ficción es invitar al lector a imaginar cómo sería su relación cognitiva con el mundo si poseyera las cualidades ilimitadas de Funes o, lo que es lo mismo, nos lleva a reflexionar acerca de la manera particularmente limitada en que nuestra memoria y nuestros sentidos finitos nos permiten conocer el mundo. Uno de los cuestionamientos más interesantes que motiva esta ficción es el de nuestras nociones de *individuo* y *clase*. Cito a Borges: «Funes no solo recordaba cada hoja de cada árbol, de cada monte, sino cada una de las veces que la había percibido o imaginado». Además, como su percepción era infalible, cada recuerdo era pleno de detalles, de manera que la siguiente vez que veía lo que para nosotros sería la *misma* hoja, a él le parecía un objeto completamente distinto: más grande o más pequeño, más arrugado o menos brillante, con límites o nervaduras diferentes, etcétera. Más o menos, probablemente, como lo que nos ocurriría a nosotros si viéramos a un niño de un año y no volviéramos a verlo hasta setenta años después: nos resultaría difícil concebirlo como el *mismo* individuo. Dice Borges sobre Funes: «No solo le costaba comprender que el símbolo genérico *perro* abarcara tantos individuos dispares de diversos tamaños y diversa forma; le molestaba que el perro de las tres y catorce (visto de perfil) tuviera el mismo nombre que el perro de las tres y cuarto (visto de frente)». Aquí se nos hace evidente el carácter *construido* de nuestra noción de *individuo*: para nosotros, que tenemos una percepción y una memoria finitas, la imagen de un perro de perfil y la del mismo perro de frente, un minuto después, son tan parecidas que sin mayor inconveniente las consideramos imágenes del *mismo* individuo; la ficción de Funes resalta el hecho de que la noción de *un mismo individuo que permanece el mismo* es una consecuencia directa de la manera finita en que percibimos y recordamos el mundo. Borges lo muestra de una manera más dramática; dice de Funes que «[s]u propia cara en el espejo, sus propias manos, lo sorprendían cada vez».

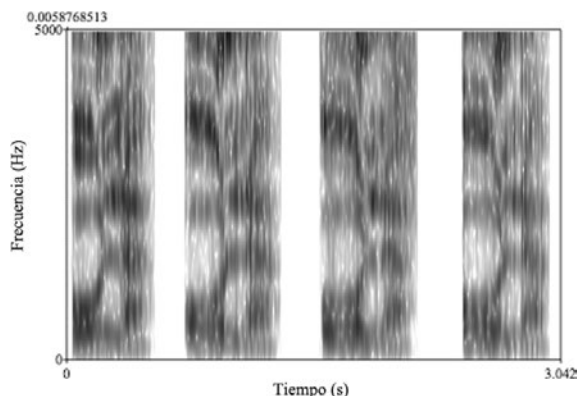
El que a Funes, como leímos líneas atrás, le costara comprender que individuos tan diversos fueran llamados «perros» muestra también el carácter *construido* de nuestra noción de *clase*. En efecto, las clases de individuos que solemos reconocer (y, para las cuales, solemos tener palabras) son constructos dependientes de nuestra percepción finita: si bien podemos reconocer diferencias entre los individuos de una misma clase, encontramos en ellos suficientes similitudes para considerarlos miembros de la misma clase (perros, gatos, sillas, etcétera). Para Funes, esto era muy difícil; según Borges: «[e]ra el solitario y lúcido espectador de un mundo multiforme, instantáneo y casi intolerablemente preciso»; y también: «[e]n el abarrotado mundo de Funes no había sino detalles, casi inmediatos». Nosotros, de manera natural, concebimos la realidad como conformada por entidades que comparten algunas propiedades y se distinguen por otras; por ello, no nos resulta extraño pensarlas agrupadas en especies, géneros o clases. Funes percibe cada entidad con tantos detalles propios que no

consigue encontrar semejanzas suficientes para concebir dos de ellas como *miembros de una misma clase*.

En un trabajito que hicimos mi colega y amigo Alfredo Arnaiz y yo hace algunos años intentamos imaginar cómo percibiría auditivamente el lenguaje humano una persona como Funes<sup>2</sup>. Dado que el funcionamiento de una lengua natural depende (entre otros muchos factores) de nuestra posibilidad de distinguir auditivamente *algunas* propiedades acústicas de los sonidos del habla pero, crucialmente, *no todas*, supusimos que Funes, al tener una percepción infalible, no podría escuchar dos estímulos acústicos distintos como sonidos iguales y, por tanto, no podría utilizar una lengua como lo hacemos nosotros. En efecto, si fuéramos como Funes y pudiéramos distinguir absolutamente todos los detalles acústicos de cada enunciado lingüístico que nos dirigen, no podríamos reconocerlo como una instancia particular de —digamos— el saludo «Hola», la pregunta «¿Cómo te llamas?» o la aseveración «Jorge Luis Borges concibe un personaje dueño de una percepción y una memoria infalibles»; cada uno de estos enunciados sería identificado como un individuo particular, único, y no como un ejemplar de una clase (la palabra «Hola», la oración «¿Cómo te llamas?», etcétera).

Examinemos las propiedades acústicas de cuatro emisiones de la palabra «Hola», producidas por la misma persona, en el espectrograma que aparece a continuación<sup>3</sup>:

Figura 1. Cuatro instancias de la palabra «Hola» producidas por el mismo hablante<sup>4</sup>



<sup>2</sup> Alfredo Arnaiz y Jorge Iván Pérez Silva, 1996. «Una lectura lingüística de “Funes el memorioso”» [manuscrito]. Lima, PUCP.

<sup>3</sup> Un espectrograma es un gráfico que representa en el eje vertical las distintas frecuencias que componen las ondas sonoras complejas que sirven como estímulos acústicos. En el eje horizontal se representa el tiempo, de modo que el espectrograma nos permite visualizar cómo cambia la estructura de frecuencias de un enunciado conforme se va articulando (ver Martínez Celdrán & Fernández Planas, 2007).

<sup>4</sup> Todos los espectrogramas han sido elaborados por el autor [nota de editores].

En la figura 1, podemos observar, representadas gráficamente como manchas en distintos tonos de gris, algunas propiedades acústicas comunes a las cuatro instancias de la palabra «Hola». Resulta claro, sin embargo, que estas cuatro emisiones no son físicamente idénticas entre sí: si bien reconocemos manchas similares en los cuatro enunciados, también comprobamos diferencias entre ellos. Ahora bien, estas diferencias acústicas no resultan ningún obstáculo para que los hablantes reconozcamos con facilidad estos cuatro ejemplares como instancias de la misma palabra. Intuyo, sin embargo, que Funes no habría podido hacer esto, pues se hubiera quedado enmarañado entre miles de detalles acústicos. Su registro de cada enunciado sería tan rico en detalles para él que sería incomparablemente distinto al de cualquier otro enunciado. Para nosotros, sencillamente, son instancias de la palabra «Hola».

En este punto, resulta conveniente preguntarse si la categorización que realizamos de distintos individuos como perros, gatos, sillas, etcétera y la correspondiente a diversos enunciados como instancias de un mismo tipo lingüístico son idénticas. Notemos, para empezar, que las primeras clases dependen notoriamente de una categorización de la realidad que involucra más consideraciones que solo la información de nuestros sentidos. Nuestra concepción del mundo resulta de un largo proceso de aprendizaje en el que desempeña un rol fundamental la interacción con otras personas: si bien nuestro conocimiento del mundo se construye a partir de la información que proporcionan nuestros sentidos, los límites conceptuales son establecidos culturalmente. A partir de la experiencia sensible, acompañada por la formación cultural, vamos construyendo los límites categoriales que dan como resultado nuestra concepción de la realidad: nuestro aprendizaje de qué individuos debemos agrupar como perros y cuáles no, qué individuos son sillas y cuáles no, depende tanto de nuestras capacidades de percepción y de retención de información sensorial, cuanto de consideraciones propias de la sociedad en la que adquirimos nuestro conocimiento del mundo.

Así, por ejemplo, a simple vista, podemos reconocer algunas semejanzas físicas entre mamíferos marinos, como las focas, los lobos marinos, las ballenas y los delfines, y diferenciarlos de mamíferos cuadrúpedos, como los caballos, las vacas y los perros. La primera de estas dos grandes clases quizá podría también subdividirse, a simple vista, en una clase compuesta por las focas y los lobos marinos, y en otra compuesta por las ballenas y los delfines aunque, ciertamente, es discutible que podamos hacerlo sin prestar especial atención a detalles o sin un entrenamiento explícito. Posiblemente, la distinción entre ballenas y delfines pueda basarse en un rasgo notorio como el tamaño, pero habrá una serie de individuos, de tamaño intermedio, que no podremos categorizar sin los límites trazados por la cultura. Y, sin duda, la distinción entre focas y lobos marinos solo podrá establecerse a partir de un aprendizaje más sofisticado.

En esa misma línea, un biólogo marino podrá realizar muchas más distinciones entre todas estas clases, gracias a su «cultura» especializada.

Esta es la gran diferencia entre Funes y nosotros. Si llevamos al extremo la ficción de Borges, tendremos que aceptar que su personaje no podía concebir clases porque su mundo estaba poblado solo por características distintivas: ningún individuo sería lo suficientemente parecido a ningún otro individuo como para considerarlos miembros de una misma clase. Es más, como vimos líneas atrás, en su mundo no habría individuos pues los cambios incesantes de propiedades se sucederían en un flujo continuo de transformaciones.

Ahora bien, para la identificación de diferentes enunciados lingüísticos como instancias de un mismo tipo —por ejemplo, las cuatro emisiones de «Hola» de la figura 1—, también es necesaria la conjunción de la percepción y la memoria con el aprendizaje cultural. La identificación de una señal acústica como un saludo, una pregunta o una aseveración (o cualquier otro acto de habla) depende crucialmente de aprender intersubjetivamente las asociaciones convencionales entre los sonidos de la lengua y las intenciones comunicativas de sus usuarios. El mismo requisito tienen que cumplir las diferentes unidades del léxico: solo podemos interpretarlas (esto es, conocer su significado) si aprendemos socialmente las convenciones que sustentan el vínculo entre su expresión y el contenido que quieren transmitir sus usuarios.

Sin embargo, hay un momento lógicamente previo al reconocimiento de estas asociaciones convencionales: primero es necesario poder identificar la expresión sonora como tal. Para dominar una lengua, el hablante-oyente debe poder reconocer en los estímulos acústicos —siempre físicamente únicos e irrepetibles— los *mismos* sonidos. La condición para esto es no distinguir todas las propiedades sonoras presentes en la señal acústica, es decir, escuchar como *igual* lo que no es físicamente igual. Esto es algo que hacemos las personas normalmente cuando aprendemos nuestra lengua materna: nuestra capacidad para distinguir sonidos del habla es mayor cuando nacemos; durante los primeros meses de vida, nuestra percepción se va adaptando a las distinciones sonoras que son pertinentes para el funcionamiento de nuestra lengua como un sistema semiótico, lo que equivale a decir que vamos *perdiendo* nuestra capacidad para distinguir sonidos del habla, pues nuestra percepción se va sintonizando al inventario cerrado de sonidos que conforma nuestro sistema fonológico (Eimas y otros, 1971; Kuhl, 1991; Kuhl y otros, 2006).

Si bien esta adaptación al medio ambiente lingüístico depende también de la interacción con otros hablantes, aprender a distinguir los sonidos del habla no es como aprender a distinguir ballenas de delfines o focas de lobos marinos. Estas categorías pertenecen a un nivel de conciencia compartido por los miembros de una comunidad, el nivel del pensamiento o conocimiento del mundo. En cambio, los sonidos

del habla pertenecen a un nivel que podríamos llamar «operativo» o «procesual»: el cerebro procesa la información acústica (siempre individualmente diferente) y la organiza en unidades (tipos sonoros), pero no con el fin de llevarlas a la conciencia —pues no es el sonido en sí mismo al que hay que atender en un intercambio lingüístico— sino como parte de las «operaciones» necesarias para la identificación de las unidades lingüísticas mayores (palabras, frases, etcétera) a las que les asociamos un significado.

Así, para reconocer diferentes instancias de la palabra *hola* es necesario poder reconocer los sonidos individuales que la conforman —[ola]—, para distinguirla, por ejemplo, de las palabras *ala*, *hora* y *ole*, cuyos sonidos individuales son, respectivamente, [ala], [ora] y [ole]<sup>5</sup>. Estas distinciones nos resultan claramente perceptibles a los hablantes de castellano, pero no lo son para todas las personas. La diferencia entre *hola* y *hora* no la notan con facilidad, por ejemplo, los hablantes de chino o japonés, pues no es parte del sistema fonológico de su lengua materna. La consecuencia de esto es que escuchan —«categorizan o clasifican auditivamente»— como un mismo sonido ambos estímulos acústicos. Como se puede ver, esta dificultad para distinguir los sonidos [r] y [l] se encuentra al otro extremo de lo que le ocurriría a Funes: mientras que los hablantes de chino o japonés categorizan auditivamente estos estímulos como un mismo sonido, Funes no podría categorizar *ningún* par de estímulos acústicos (por más parecidos que fueran para nosotros) como instancias de un mismo sonido —digamos, [r]— porque todo estímulo acústico sería para él completamente distinto a otro.

Así, pues, de acuerdo con nuestra percepción y memoria limitadas, y en función de los estímulos acústicos propios de nuestro medio ambiente lingüístico, los hablantes-oyentes aprendemos a procesar auditivamente el habla. Propongo que esta organización particular de la materia acústica que realiza nuestro cerebro de acuerdo con su estructura propia y a partir de los estímulos lingüísticos a los que ha estado expuesto desde el nacimiento —organización que determina la forma en que escuchamos— genera *unidades perceptuales* que se distinguen de las ondas sonoras mismas, de la respuesta anatómico-fisiológica del oído, así como de las operaciones físicas del cerebro, de las que emergen. Estas unidades de percepción se distinguen, a su vez, de las *unidades conceptuales* que emergen de la organización de la información que realiza nuestro cerebro cuando categorizamos individuos en clases como perro, gato o silla, lo que, como vimos, resulta en nuestro conocimiento del mundo.

---

<sup>5</sup> La representación gráfica de los sonidos individuales como letras entre corchetes es la convención más usual en las ciencias del lenguaje. Las letras pertenecen al Alfabeto Fonético Internacional propuesto por la Asociación de Fonética Internacional (1999).

El hecho de que podamos distinguir entidades y procesos físicos probadamente diferentes (estímulos acústicos) que corresponden a una sola unidad perceptual (sonido) resultante del funcionamiento de nuestro cerebro es razón suficiente para describir a aquellos con un lenguaje fisicalista y a esta con un lenguaje mentalista: algo como «dos estímulos acústicos de tales y tales frecuencias son procesados por el cerebro [lenguaje fisicalista] y percibidos como un solo sonido del habla [lenguaje mentalista]». Esto no significa comprometerse con una ontología dualista. Equivale simplemente a reconocer el hecho de que las unidades perceptuales (como las conceptuales) no son reductibles a las propiedades acústicas de las ondas sonoras, sino que resultan de la especial organización que realiza nuestro cerebro de estas<sup>6</sup>.

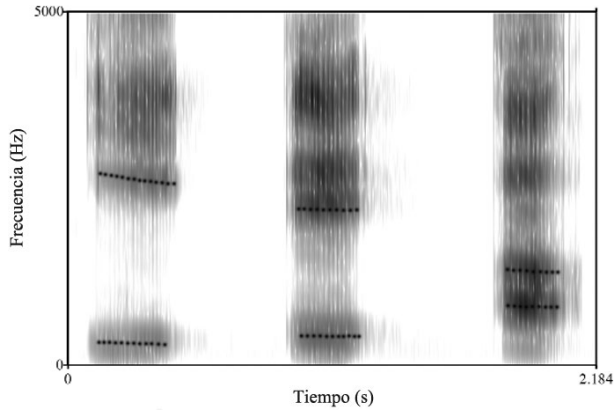
Ahora bien, desde los orígenes de la lingüística y la fonología modernas, con De Saussure (1945[1916]), Trubetzkoy (1968[1939]) y Jakobson (1968[1941]), se ha subrayado la importancia fundamental que tienen para el funcionamiento de los sistemas lingüísticos, las *diferencias* entre sus unidades. Para De Saussure, en efecto, cada lengua ofrece una organización particular de la materia sonora para la constitución de los signos que utilizan sus usuarios: «Cada idioma compone sus palabras a base de un sistema de elementos sonoros, cada uno de los cuales forma una unidad netamente deslindada y cuyo número está perfectamente determinado. Pero lo que los caracteriza no es, como se podría creer, su cualidad propia y positiva, sino simplemente el hecho de que no se confunden unos con otros. Los fonemas son ante todo entidades opositivas, relativas y negativas» (1945[1916], p. 201).

El caso de los sonidos vocálicos es particularmente ilustrativo de cómo los sonidos del habla no se definen positivamente sino por oposición. Los sonidos vocálicos se producen modificando, por resonancia, la onda compleja de las cuerdas vocales mediante variaciones en el tamaño de las cavidades faríngea y bucal. La siguiente figura muestra espectrogramas de los sonidos vocálicos [i, e, a] del castellano:

---

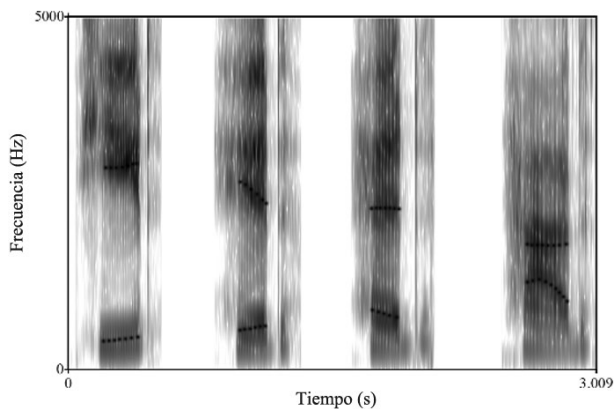
<sup>6</sup> El recurso a nociones mentalistas para describir la manera en que nuestro cerebro organiza la información del medio ambiente se encuentra en los orígenes mismos de la ciencia cognitiva contemporánea (ver Thagard, 2008; Westbury & Wilensky, 2001). La gramática generativa, gestada desde finales de la década de 1950 por Noam Chomsky, se sustenta en la idea de que el cerebro humano organiza la información lingüística mediante representaciones abstractas que corresponden —pero, crucialmente, no se identifican con— los estímulos acústicos que les sirven de expresión. Estas representaciones abstractas —de carácter esencialmente operativo o procesual— son de distintos tipos, de acuerdo con el nivel de organización o descripción del sistema lingüístico al que pertenecen: representaciones fonológicas —como las unidades perceptuales de las que hemos venido hablando— morfológicas, sintácticas, etcétera. Ver Chomsky (1986) y Jackendoff (2002), quien engloba los diferentes niveles de procesamiento lingüístico bajo el nombre de nivel «mental-f(uncional)».

Figura 2. Algunas vocales del castellano



Si bien los espectrogramas presentan semejanzas, hay una diferencia acústica que determina que estos estímulos acústicos sean percibidos como vocales diferentes por los hablantes de castellano: las frecuencias de las dos franjas horizontales oscuras inferiores (marcadas con puntos para apreciarlas mejor). Las medidas de estas dos franjas, conocidas como «formante 1» y «formante 2», son, de acuerdo con los estudiosos, las claves acústicas que determinan el timbre o calidad de los sonidos vocálicos. Veamos la siguiente figura, que muestra los espectrogramas de las palabras *heed*, *hid*, *head* y *had* del inglés (los formantes 1 y 2 de las vocales respectivas aparecen marcados con puntos):

Figura 3. Algunas vocales del inglés



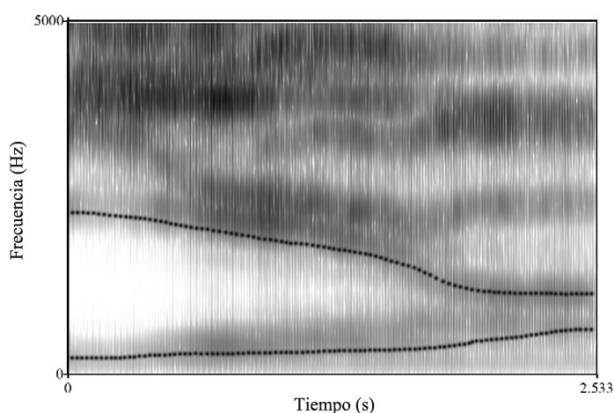
Como en el caso anterior, estos espectrogramas también presentan los formantes 1 y 2 en diferentes posiciones, lo que determina que los estímulos sean percibidos

como sonidos diferentes. Ahora bien, ocurre que los hablantes de castellano con poco entrenamiento en el inglés no distinguimos normalmente todos los sonidos que sí distinguen sin dificultad los anglohablantes. Es conocido el hecho de que solemos escuchar la palabra *hid* bien como *heed*, bien como *head*, y que escuchamos una palabra como *hood* con la misma vocal que *who'd*.

Así, pues, tanto el inglés como el castellano utilizan un número finito de sonidos vocálicos para la construcción de sus signos, pero los sonidos no se corresponden unos con otros y no pueden definirse en términos acústicos independientemente del uso que se les da en cada lengua. Más bien, los sonidos lingüísticos se definen por *oposición* entre ellos dentro de cada lengua porque cada uno corresponde a un rango acústico cuyos límites están determinados por los rangos de los sonidos vecinos.

Esto se puede ilustrar con el siguiente espectrograma, que corresponde a un continuo sonoro que va desde un estímulo acústico perceptible como [i] (al inicio) hasta uno perceptible como [a] (al final):

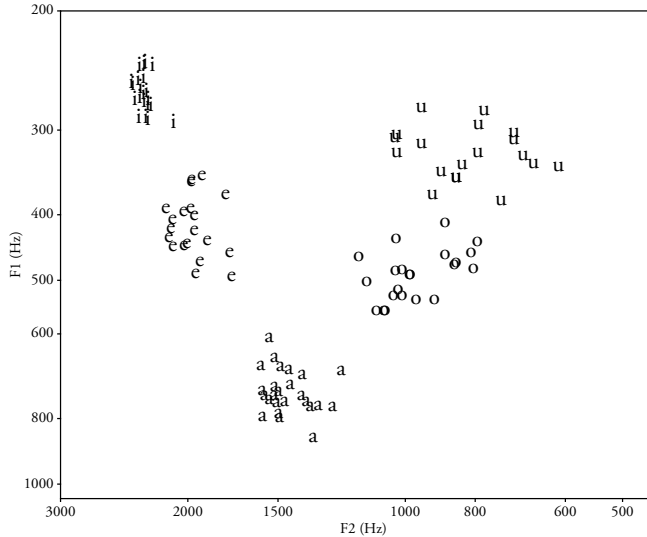
Figura 4. Continuo acústico



La figura 4 muestra cómo los formantes 1 y 2 van acercándose paulatinamente. Los estímulos que se perciben de manera similar a [i] presentan mayor separación entre sus formantes y los que se perciben de manera similar a [a] presentan formantes muy cercanos. Entre estos dos extremos, los estímulos acústicos se perciben de acuerdo con la lengua que uno haya aprendido, dado que, a lo largo de este continuo, diferentes lenguas delimitan distintas unidades sonoras: el inglés distingue, al menos, las cuatro vocales de la figura 3 (si no más) y el castellano distingue las tres de la figura 2.

Lo más interesante es que las articulaciones particulares de cada sonido no son idénticas, sino que se dispersan en un campo que cubre gran parte del continuo, como muestra la siguiente figura:

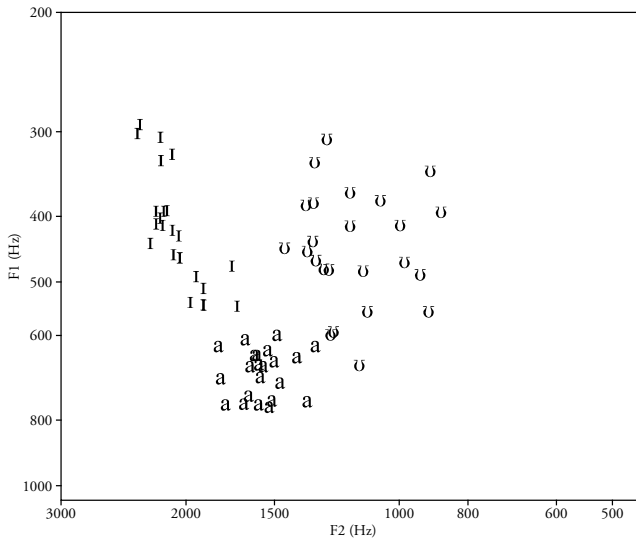
Figura 5. Dispersión de los estímulos acústicos correspondientes a las vocales castellanas



Fuente: Pérez, Acurio y Bendezú (2008, p. 15).

Así, la [i] y la [e] del castellano cubren más o menos el mismo espacio que cubren la [i, ɪ, e, æ] del inglés. Por otra parte, la unidad perceptual [ɪ] del quechua cubre el espacio de la [i] y la [e] del castellano, como se ve en la siguiente figura:

Figura 6. Dispersión de los estímulos acústicos correspondientes a las vocales quechuas



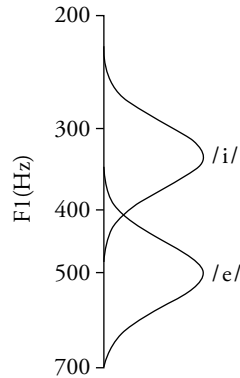
Fuente: Pérez, Acurio y Bendezú (2008, p. 18).

Así, pues, dentro de cada lengua los hablantes suelen producir diferentes estímulos acústicos con la intención de producir una misma unidad perceptual y estos distintos estímulos son, de hecho, escuchados como instancias de un solo sonido por los hablantes de la misma lengua. Las distintas instancias de las vocales inglesas de *heed* o de *hid* no son nunca idénticas acústicamente: lo que las identifica no son sus propiedades acústicas inherentes sino su relación de oposición con sus vocales vecinas. Y lo mismo ocurre con las vocales del castellano: como muestra la figura 5, escuchamos como /e/ los estímulos que no sean tan altos que se confundan con /i/ ni tan bajos que se confundan con /a/; pero no podemos definirlos en términos acústicos absolutos. Esta característica de las lenguas naturales haría imposible que Funes escuchara los sonidos del habla como nosotros: no podría escuchar como auditivamente igual lo que es acústicamente diferente. Para nosotros, en cambio, gracias a nuestras capacidades perceptivas limitadas, resulta fácil el agrupamiento de estímulos acústicos distintos en categorías auditivas<sup>7</sup>.

Los estudiosos han mostrado que existe una relación directa entre la manera en que los hablantes producen sus vocales en una comunidad de habla y la manera en que aprenden a percibirlos. En palabras de Escudero, «la experiencia con las formas concretas con que se producen los sonidos resulta en una percepción óptima a partir de la cual los oyentes aprenderán a categorizar la señal hablada en vocales y consonantes que coincidirán con lo producido por el hablante» (2005, p. 52; la traducción es mía). Como vimos, los estímulos acústicos que corresponden a una unidad perceptual no son acústicamente idénticos. El siguiente gráfico muestra, con respecto a la medida del formante 1, que la mayoría de estímulos acústicos que sirven como realización material del sonido [i] en castellano se encuentran alrededor de los 350 Hz, pero que también hay un gran porcentaje de estímulos que se encuentran por encima y por debajo de esta medida.

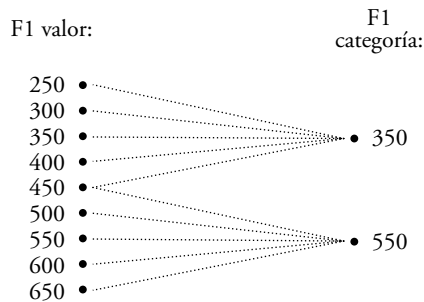
---

<sup>7</sup> En Arnaiz y Pérez (1996) la tesis era que Funes no podría dominar una lengua. Eduardo García Ramírez (comunicación personal), sin embargo, señala correctamente que las habilidades extraordinarias de Funes sí le permitirían calcular los límites de variación correspondientes a las diferentes unidades lingüísticas de una lengua y, por tanto, convertirse en un hablante-oyente de la misma. De todas maneras, nuestra forma de aprender una lengua sería diferente de la de Funes, en tanto nosotros identificamos unidades perceptuales por hipodiferenciación, mientras él lo haría por memorización de los distintos estímulos acústicos que los hablantes usamos de manera equivalente. Agradezco a Eduardo García Ramírez su atinado comentario.

**Figura 7. Promedios de producción y distribuciones de [i] y [e] del castellano**

Fuente: Escudero (2005, p. 53).

Lo mismo se puede apreciar con respecto al sonido [e], pero en relación con la medida de 550 Hz. La siguiente figura representa la manera en que los oyentes construyen sus categorías auditivas sobre la base de agrupaciones o «mapeos» de estímulos acústicos diferentes en un solo sonido que se distingue de manera discreta o categórica de otro.

**Figura 8. Mapeo del continuo acústico a las categorías auditivas discretas**

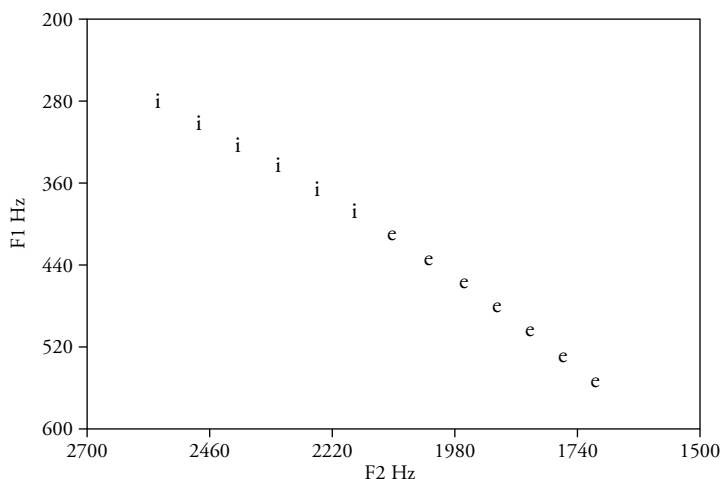
Fuente: Adaptado de Escudero (2005, p. 45).

Así, pues, el resultado final del aprendizaje de una lengua es un estado determinado del cerebro en función del cual procesamos u organizamos los estímulos acústicos. Como consecuencia de esto, los cerebros de hablantes de diferentes lenguas procesan los mismos estímulos acústicos de maneras diferentes. Una metáfora usual de este fenómeno es la del «filtro»: se dice que los hablantes «filtran» los estímulos acústicos a través de las categorías de su sistema fonológico.

Un ejemplo notable de esto es la distinta manera en que los hablantes de castellano y los de quechua percibimos los sonidos vocálicos. Napurí (2012) muestra

que ante un continuo de estímulos acústicos sintetizados, los hablantes nativos de castellano distinguimos las vocales [i] y [e] de manera categórica, de acuerdo con la siguiente figura:

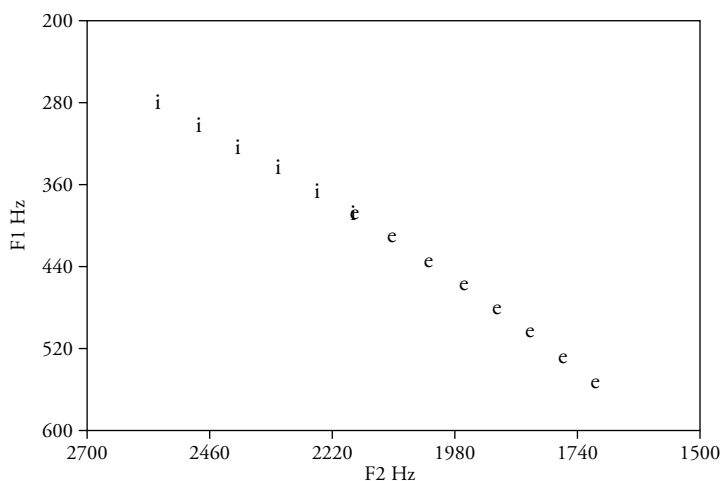
**Figura 9. Percepción de hablantes nativos de castellano 1**



Fuente: Napurí (2012, p. 56).

Como se ve, algunos hablantes oyentes escuchan los seis primeros estímulos como [i] y los siguientes siete como [e]; otros, en cambio, categorizan el sexto estímulo a veces como [i] y a veces como [e]:

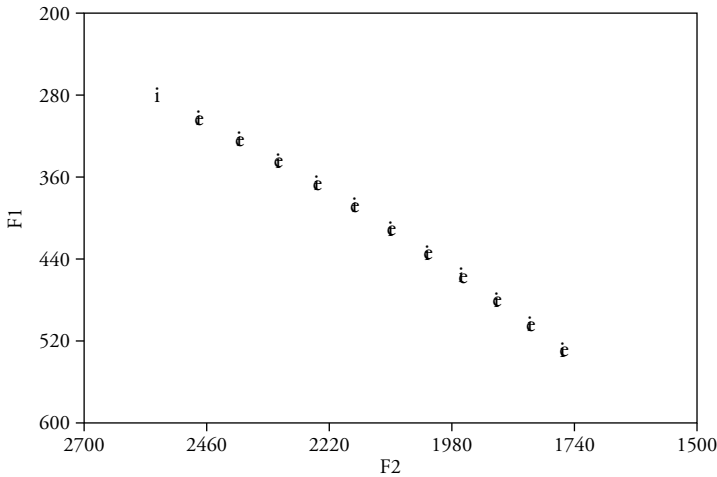
**Figura 10. Percepción de hablantes nativos de castellano 2**



Fuente: Napurí (2012, p. 57).

Por su parte, los bilingües quechuahablantes que tienen un conocimiento muy incipiente de castellano identifican solo el primer estímulo como [i], mientras que los otros doce los categorizan a veces como [i] y a veces como [e]:

Figura 11. Percepción de bilingües incipientes quechua-castellano

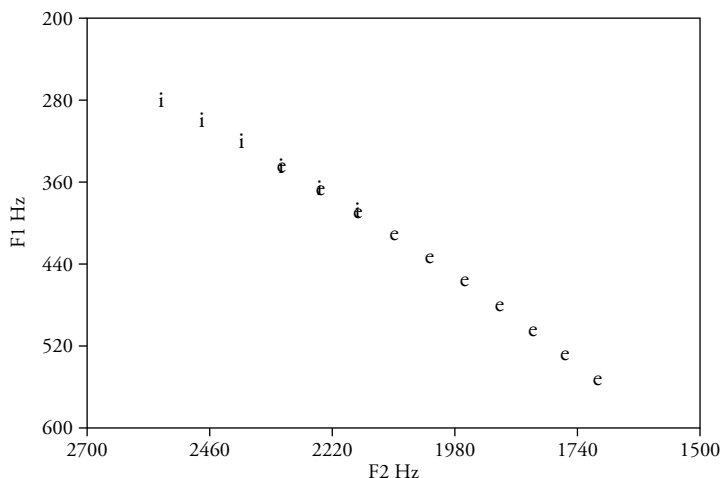


Fuente: Napurí (2012, p. 30).

Esta identificación aleatoria demuestra, de acuerdo con Napurí (2012), que la percepción de los estímulos está siendo «filtrada» por las categorías fonológicas de la lengua materna dominante de los bilingües quechuahablantes; en efecto, como vimos antes, el quechua solo distingue una unidad perceptual allí donde el castellano distingue dos.

Algo que resulta particularmente interesante es que, conforme los bilingües van aprendiendo más y más castellano, van distinguiendo cada vez más los estímulos acústicos. Si bien Napurí (2012) no realiza un estudio longitudinal sobre el desarrollo de la percepción de los bilingües, su investigación transversal, con distintos individuos en diversos niveles de aprendizaje del castellano, nos da una idea clara sobre cómo van construyendo y modificando sus categorías auditivas, hasta que los bilingües avanzados muestran una categorización idéntica a la de los monolingües. La figura siguiente presenta la categorización de un bilingüe quechua-castellano intermedio:

**Figura 12. Percepción de bilingües intermedios quechua-castellano**



Fuente: Napurí (2012, p. 50).

En conclusión, la manera en que el cerebro humano procesa la información lingüística sugiere la emergencia de unidades perceptuales o auditivas que claramente no son idénticas a sus correlatos físicos. La evidencia empírica, presentada en este trabajo, sobre la distinta manera en que los hablantes de quechua, castellano e inglés perciben los mismos estímulos acústicos muestra la conveniencia de describir los sistemas fonológicos en términos mentalistas, en la medida en que el funcionamiento de los sistemas lingüísticos —en particular, la distinción de palabras— no depende de los estímulos acústicos en sí mismos sino en tanto organizados lingüísticamente por el cerebro en tales unidades perceptuales<sup>8</sup>.

<sup>8</sup> Estas unidades perceptuales no son las únicas unidades fonológicas que presentan carácter mental. La sutil organización de los sistemas fonológicos incluye unidades menores a las unidades perceptuales —rasgos segmentales y suprasegmentales—, así como unidades mayores a ellas —sílabas, pies, patrones acentuales, entonacionales, etcétera—. A esto habría que añadir la existencia de reglas o restricciones fonológicas, dependiendo del paradigma teórico asumido. Ver Goldsmith (1994).

## REFERENCIAS BIBLIOGRÁFICAS

- Asociación de Fonética Internacional (1999). *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press.
- Borges, Jorge Luis (1983). *Narraciones*. Bogotá: Oveja Negra.
- Chomsky, Noam (1986). *Knowledge of Language: Its Nature, Origin, and Use*. Nueva York: Praeger.
- De Saussure, Ferdinand (1945[1916]). *Curso de lingüística general*. Buenos Aires: Losada.
- Eimas, Peter D., Einar R. Siqueland, Peter Jusczyk & James Vigorito (1971). Speech Perception in Infants. *Science*, 171, 971-974.
- Escudero, Paola (2005). *Linguistic Perception and Second Language Acquisition*. Utrecht: LOT.
- Goldsmith, John (ed.) (1994). *The Handbook of Phonological Theory*. Cambridge: Blackwell.
- Jackendoff, Ray (2002). *Foundations of Language*. Oxford: Oxford University Press.
- Jakobson, Roman (1968[1941]). *Child Language, Aphasia and Phonological Universals*. La Haya: Mouton.
- Kuhl, Patricia K. (1991). Human Adults and Human Infants Show a «Perceptual Magnet Effect» for the Prototypes of Speech Categories, Monkeys Do Not. *Perception and Psychophysics*, 50, 93-107.
- Kuhl, Patricia K. y otros (2006). Infants Show a Facilitation Effect for Native Language Phonetic Perception Between 6 and 12 Months. *Developmental Science*, 9(2), F13-F21.
- Martínez Celdrán, Eugenio & Ana María Fernández Planas (2007). *Manual de fonética española*. Barcelona: Ariel.
- Napurí, Andrés (2012). *Adquisición de la percepción de vocales del español por bilingües quechua-español*. Tesis de maestría, PUCP.
- Pérez, Jorge, Jorge Acurio & Raúl Bendezú (2008). *Contra el prejuicio lingüístico de la motosidad*. Lima: Instituto Riva Agüero-PUCP.
- Thagard, Paul (2008). *La mente: introducción a las ciencias cognitivas*. Buenos Aires: Katz.
- Trubetzkoy, Nikolai (1968[1939]). *Principios de fonología*. Buenos Aires: Kapelus.
- Westbury, Chris & Uri Wilensky (2001). La representación del conocimiento en la ciencia cognitiva. Sus implicancias en la educación. En Ministerio de Educación del Perú-MINEDU, *¿Sabemos cómo se aprende? Nuevos enfoques sobre el aprendizaje* (pp. 25-87). Lima: MINEDU.