

Proyecto de investigación

Consonantes oclusivas y el efecto McGurk: una explicación fonético-acústico-perceptiva

Silvana Andrea D'Ottone Campana

Profesor guía: Domingo Román

1. Introducción

El efecto McGurk es una ilusión perceptiva audiovisual de efecto fuerte. En éste, la información acústica (audio) de un fonema es incongruente con la información articulatoria (video) presentada al espectador. Al intentar procesar ambas modalidades, la persona tiende a combinar o fusionar las informaciones de los dos estímulos. El ejemplo más común es presentar a alguien el video de la articulación de la sílaba [ga] pero con el audio sustituido por el de [ba]. El resultado debería ser que el oyente-espectador percibiera [da]; es decir, ninguno de los fonemas que están siendo presentados, sino uno nuevo. Los autores han denominado esta ilusión como 'fusión'.

Tanto en ésta como en otras combinaciones, las respuestas de los oyentes-espectadores dependerán de cómo integren las distintas informaciones visuales y auditivas que perciban simultáneamente. Es decir, las respuestas estarán determinadas por los rasgos acústicos y articulatorios (al menos lo superficialmente perceptible) de las informaciones incongruentes y cómo éstos son procesados por quien ve y escucha.

Dicho de otro modo, los rasgos acústicos de las consonantes, por un lado, y las pistas visuales de la articulación de las consonantes, por otro lado, forman parte de la respuesta perceptiva de la persona expuesta al efecto McGurk. La pregunta es, ¿Cuáles de dichos rasgos son determinantes para la percepción de un fonema?

Esta investigación nace como un intento de aplicar y verificar la teoría que existe en relación al efecto McGurk en un contexto de habla hispana, específicamente, del español de Chile, dada la limitada cantidad de información existente. Además, en vista de las explicaciones que se han tratado de hacer en otros ámbitos de conocimiento, parece necesaria la participación de la lingüística, ya que puede ser un aporte importante para comprender el fenómeno. Cabe señalar, además, la relevancia que tendría el obtener un mayor conocimiento del efecto McGurk, debido a los distintos ámbitos en los que puede ser aplicado (sicología, salud, enseñanza de lenguas). Esto justifica la necesidad de explorarlo desde sus distintos componentes; en este caso, la perspectiva del estudio será principalmente fonética.

A la vez que este estudio intenta entregar datos que puedan contribuir con la explicación del fenómeno mencionado, también parte de una carencia en el ámbito de la fonética de nuestro país en el área de la percepción. La presente investigación, por tanto, podría contribuir en la descripción de los rasgos determinantes para la percepción de los fonemas oclusivos del español de Chile.

2. Preguntas de investigación

¿Por qué ante una realización incongruente (McGurk) los sujetos perciben, mayoritariamente, un fonema y no otro?

¿Qué marcas o rasgos acústicos y articulatorios son determinantes para la percepción de un fonema consonántico oclusivo ?

3. Marco Teórico

3.1 La percepción del habla:

Breve introducción al tema y a las distintas teorías o modelos de percepción del habla. La base perceptual y la lengua materna (Gòsy, 1992)

3.1.1 La teoría motora (Gòsy, 1992; Thatam & Morton, 2006)

3.1.2 Modelo de análisis por síntesis (Gòsy, 1992; Thatam & Morton, 2006)

3.1.3 Modelo de percepción global del habla (Gòsy, 1992)

3.1.4 Modelos híbridos de percepción del habla: categorías, realismo directo, teoría de los cuanta – la búsqueda de la invariancia (Thatam & Morton, 2006)

3.2 Las consonantes oclusivas:

3.2.1 Descripción general (Quilis, 1981; Ladefoged, 2005)

3.2.2 Las consonantes oclusivas del español (Quilis, 1981; Martínez Celdrán & Planas, 2007)

3.2.3 Las consonantes oclusivas del español de Chile (Soto Barba, 1994; Poblete, 1995; Soto Barba y Roldán, 1997; Pérez, 1998; Soto Barba y Valdivieso, 1999)

3.3 El efecto McGurk:

3.3.1 Breve reseña histórica y explicación (Colin *et al*, 2002; Behne *et al*, 2007; Hardison *et al*, 2007)

3.3.2 El efecto McGurk desde la psicología y la neurociencia (Nichols *et al*, 2003; Gentilucci *et al*, 2005; Omata *et al*, 2008; Baart *et al*, 2010)

3.3.3 Panorámica del desarrollo del efecto McGurk (especialmente en trastornos del lenguaje o personas con impedimento perceptivo)

3.4 Teoría de la optimalidad (Boersma, 2006)

4. Objetivos

Contribuir a la descripción perceptiva de los fonemas oclusivos (qué rasgos son relevantes para el oyente).

Describir el efecto McGurk en el español de Chile.

Explicar las respuestas ante el efecto McGurk en base a los rasgos acústicos y articulatorios de las consonantes presentadas

Comparar los resultados con los obtenidos en experimentos realizados en otras lenguas.

Discutir y contrastar los datos obtenidos en relación a los estudios fonéticos existentes de los fonemas oclusivos del español de Chile.

5. Hipótesis

El tipo de respuesta (auditiva, ilusión, combinación) percibida por el oyente-espectador, está relacionada con el tipo de estímulo presentado (qué tipo de consonantes han sido combinadas y en qué modalidades).

6. Metodología

6.1 Participantes:

90 sujetos (18-25 años) que no tengan algún tipo de impedimento auditivo o visual (en este último caso, si se trata de un problema corregido, el sujeto será igualmente válido para el procedimiento).

6.2 Procedimiento:

El experimento consiste en la presentación de videos con audio, incongruentes y congruentes, de sílabas consonánticas sonoras y sordas del español de Chile: /ba-da-ga/ /pa-ta-ka/. La realización de los videos se hará con una cámara (x) y con un editor de videos (Windows Movie Maker) El experimento constará de 24 estímulos¹:

baVbaA - baV daA - baV gaA - baVpaA
daV daA - daV baA - daV gaA - daVtaA
gaVgaA - gaV baA - gaV daA - gaVkaA

paVpaA - paV taA - paV kaA - paVbaA
taVtaA- taV paA - taV kaA - taVdaA
kaVkaA- kaV paA - kaV taA - kaVgaA

¹ La A mayúscula corresponde al estímulo de audio, mientras que la V mayúscula indica qué articulación de sílaba es presentada en el video.

Cada sujeto tendrá a su disposición un equipo con audífonos y video. Además, se le entregará una hoja de respuestas en la cual deberá marcar la sílaba que cree haber percibido después de ver y oír cada uno de los 24 videos. A continuación, se presenta un ejemplo de las opciones entre las cuales podrá elegir el espectador:

En el caso de que se presente una sílaba 'ba' en la modalidad visual, junto con una 'da' auditiva

1. ba
2. da
3. bda
4. dba
5. ga
6. pa

La primera opción corresponde al estímulo visual. En este caso, se considera que se trata de una ilusión, ya que la persona incorpora el movimiento de los labios como información articulatoria, y la información acústica, la cual le informa de la sonoridad.

La segunda opción corresponde sólo a la información acústica. En este caso, se estima que la persona no incorpora, o bien, ignora la información articulatoria, por lo cual no se puede considerar como un proceso audiovisual.

Las opciones tres y cuatro, 'bda' y 'dba', corresponden a combinaciones entre las informaciones acústica y articulatoria. Es decir, el sujeto incorpora ambos estímulos en la percepción.

La opción 'ga' es lo que los estudios del efecto McGurk han denominado 'fusión'. En este caso, los individuos procesan ambas modalidades, acústica y visual, y perciben un fonema que no corresponde a ninguno de los estímulos presentados.

La última opción contiene la consonante correlativa de la serie sorda. Si el sujeto percibe esta sílaba, se considera que no ha habido una integración audiovisual, sino sólo visual.

7. Resultados

Los datos serán analizados mediante la prueba estadística de chi-cuadrado, la cual pone a prueba la hipótesis de nulidad (H0) según la cual dos variables no están relacionadas; es decir, las variables observadas son independientes. En el caso de esta investigación, se intentará determinar si existe una relación de dependencia entre las variables 'tipo de respuesta' y 'tipo de estímulo presentado'. Si la significación estadística de la prueba es suficiente como para asumir una relación de dependencia de las variables, se considerará que éstas son dependientes (están relacionadas). En este caso, sin embargo, la prueba de chi-cuadrado sólo indica si la diferencia entre las frecuencias esperadas y las observadas son lo suficientemente grandes como para rechazar una fluctuación aleatoria, pero no nos indica dónde están esas diferencias. Por esta razón, el análisis debe operar en dos niveles: uno descriptivo, en donde se observan los datos de manera 'artesanal', y uno inferencial, en el cual se recurre al valor arrojado por el análisis correspondiente a los “residuos tipificados corregidos”.

Para trabajar los datos como variables, se reunirán las respuestas en tres grandes tipos: auditiva, combinación e ilusión. La primera corresponde a la percepción del estímulo auditivo. En la segunda se agrupan las opciones que involucran ambos estímulos. La ilusión, corresponderá ya sea a una respuesta de fusión o a la respuesta visual². Esta primera variable se denominará 'tipo de respuesta'.

La segunda variable corresponderá al tipo de estímulo creado; es decir, qué tipos de consonantes han sido utilizados, y en qué modalidades, en el estímulo presentado. En este caso, las categorías son:

bA dV	(consonante bilabial Audio + consonante dental Video)
bV vV	(consonante bilabial Video + consonante velar Video)
dA bV	(consonante dental Audio + consonante bilabial Video)
dV vA	(consonante dental Video + consonante velar Audio)
vA dV	(consonante velar Audio + consonante dental Video)
vV bA	(consonante velar Video + consonante bilabial Audio)

En la prueba de chi-cuadrado, se verificará la hipótesis nula (las variables no están relacionadas)

2 Se toman en cuenta las respuestas visuales que no presentan confusión de sonoridad (por ejemplo, confundir 'b' con 'p'), puesto que ellas implican necesariamente la incorporación de la lectura de labios y la información auditiva, la cual da la información de sonoridad (Behne et al, 2002).

entre 'tipo de respuesta' y 'tipo de estímulo'. Si ésta no se comprueba, se asumirá que existe algún tipo de relación entre la forma en que se combinan las consonantes en el estímulo creado, y el efecto que esto tiene en la percepción del mismo por parte del espectador.

Además, para determinar si la sonoridad del estímulo presentado tiene algún efecto en la percepción, se añadirá la variable 'sonoridad' como capa en una segunda prueba de chi-cuadrado.

Bibliografía:

1. Efecto McGurk:

Baart, Martijn y Jean Vroomen. (2010) "Do you see what you are hearing? Cross-modal effects of speech sounds on lipreading" *Neuroscience Letters* 471,100–103

Behne, D., Wang, Y. Alm, M., Arntsen, I., Eg, R., & Valsø, A. (2007). "Changes in audio-visual speech perception during adulthood". *International Conference on Auditory-Visual Speech Processing*

Campbell, Ruth (1992) "The neuropsychology of lipreading". *Biological Sciences* 335, 39-45

Colin, C., Radeau, M., Deltenre, P., & Demolin, D. (2002) "The role of sound intensity and stop-consonant voicing on McGurk fusions and combinatios". *European Journal of Cognitive Psychology* 475 – 491

Gentilucci Maurizio y Luigi Cattaneo (2005) "Automatic audiovisual integration in speech perception" *Exp Brain Research* 167, 66–75

Hardison D. (2007). "The visual element in phonological perception and learning". En M. C. Pennington (Ed.), *Phonology in context* (pp. 135-158). New York: Palgrave Macmillan.

Massaro, Dominic and Michael m. Cohen (1990) "Perception of synthesized audible and visible speech" *Psychological Science*, 1

Nicholls, Michael, Dara A. Searle, and John L. Bradshaw. (2003) "Read My Lips: Asymmetries in the Visual Expression and Perception of Speech Revealed Through the McGurk Effect". *Psychological Science* 15, 138-141 2003

Omata, Kei and Ken Mogi (2008) "Fusion and combination in audio-visual integration" *Proceedings of the Royal Society* 464, 319–340

Windmann, Sabine (2004) "Effects of sentence context and expectation on the McGurk illusion". *Journal of Memory and Language* 50, 212-230

2 . Oclusivas:

Pérez, Hernán (1998) "Incidencia de dos rasgos acústicos en la percepción de la correlación /p, t, k/ vs. /b,d,g/". *Revista de lingüística teórica y aplicada* 36, 113 – 126

Poblete B., María Teresa (1995c). "La sonorización de las obstruyentes sordas /p,t,k/ en el habla urbana de Valdivia". *Estudios filológicos* 27, 73 – 94

Roldán, Yasna y Jaime Soto Barba (1997). "El VOT de /p,t,k/ y /b,d,g/ en el español de Valdivia: un análisis acústico".

Estudios filológicos 32, 27-33

Soto Barba, Jaime (1994). “¿Los fonemas /b/ y /p/ se diferencian por la sonoridad?” *Estudios filológicos* 29, 33 – 38

Soto Barba, Jaime y Humberto Valdivieso (1999). “Caracterización fonético-acústica de la serie de consonantes /p,t,k/ vs. /b,d,g/” *Onomázein* 4, 125-133

Vivanco, Hiram (1999). “The VOT of bilabial stops in English and in Spanish”. *Phonetics group bulletin*, 8, 20-32

3. Libros:

Gósy, Mária (1992) *Speech Perception*. Frankfurt: Hector

Lieberman, Philip y Sheila Blumstein (1988) *Speech physiology, speech perception, and acoustic phonetics*. Cambridge: Cambridge University Press

Martínez, Eugenio y Ana M. Fernández. (2007) *Manual de fonética española*. Barcelona: Ariel

Tatham, Mark y Katherine Morton. (2006) *Speech production and perception*. Hampshire: Palgrave Macmillan

4. Teoría de la optimalidad

Boersma, Paul (2006) “A constraint – based explanation of the McGurk effect”.

Prince, Alan & Paul Smolensky (2002) *Optimality Theory: Constraint interaction in Generative Grammar*.