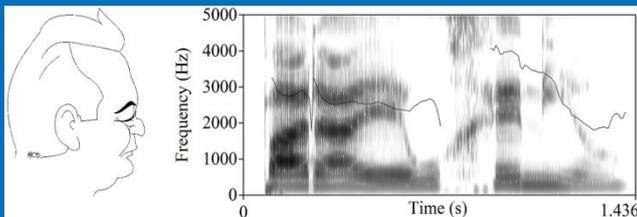


Gesto y voz en la percepción de enunciados declarativos e interrogativos del español

Hernán Martínez

Elsa Mora



Fernández Planas, A. Ma. (ed.) (2016): *53 reflexiones sobre aspectos de la fonética y otros temas de lingüística*, Barcelona, págs. 243-252.

ISBN: 978-84-608-9830-6.

Gesto y voz en la percepción de enunciados declarativos e interrogativos del español

Hernán Martínez
Universidad de Los Andes
hernan.martinez@ula.ve

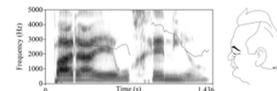
Elsa Mora
Universidad de Los Andes
elsamora@ula.ve

Para Eugenio en agradecimiento a sus enseñanzas, sugerencias y oportunidades brindadas a los miembros del Laboratorio de Fonética de la Universidad de Los Andes, Mérida-Venezuela.

1. INTRODUCCIÓN

Además de marcas sintácticas y entonativas, pareciera ser que los usuarios de las lenguas utilizan marcas gestuales para generar la diferencia entre enunciados declarativos e interrogativos. Las expresiones faciales, los movimientos de la cabeza y la mirada, parecen estar profundamente relacionados con la producción y la percepción del habla. Este hecho lo han señalado Borrás, Kaland, Prieto y Swerts (2013) en un estudio en el que revelan que tanto en la producción como en la percepción de enunciados interrogativos en holandés y catalán existen diferencias auditivas específicas que los hacen distintos de otro tipo de enunciado, pero también señalan que la mirada y el movimiento de las cejas son marcas gestuales que se utilizan para crear esta distinción. Estas afirmaciones también son hechas para el catalán por Crespo, Kaland, Swerts y Prieto (2013). Al analizar la producción de enunciados interrogativos en alemán y catalán, estos autores también afirman que en alemán las marcas entonacionales son más importantes que las gestuales. En todo caso, en estos dos trabajos se afirma que la percepción depende en gran medida de la información auditiva para determinar si un enunciado es una pregunta o una afirmación, pero la precisión se mejora aún más cuando se añade información visual.

La influencia de la información visual durante la percepción de enunciados declarativos e interrogativos también fue estudiada en el español venezolano por Higuera (2013). En su trabajo, la autora sostiene que durante el proceso perceptivo la información visual juega un papel preponderante y decisivo. Afirma, asimismo, que los oyentes toman sus decisiones perceptivas mayormente con base en lo que ven y no tanto en lo que escuchan. Al igual que en otras investigaciones ya citadas, Higuera señala que cuando los oyentes cuentan con un input visual y uno auditivo incongruentes, los oyentes presentan dificultades para percibir el mensaje, decidiendo en la mayoría de los casos por la información visual. Este hecho también ha sido señalado por Martínez, Rojas y Suárez (2012), quienes señalan que los oyentes están cotejando siempre ambos inputs durante el



proceso de percepción del habla. Estos autores confirman que la atención recae en aquel input que mejor se adapte a las expectativas (contextuales/situacionales, culturales, sociales, anímicas, etc.) del oyente. También señalan que los oyentes atienden mejor al input visual que al auditivo durante la percepción audiovisual. Tanto Higuera (2013) como Martínez, Rojas y Suárez (2012), quienes han trabajado con estímulos manipulados a los que se les ha sustituido el audio o el video, afirman que el input visual es importante a la hora de escoger ante estímulos manipulados o incongruentes: cuando se presentan este tipo de estímulos, en los que el input visual no se corresponde con el input auditivo, los oyentes seleccionan mayoritariamente la información visual.

Partiendo de la observación que hemos hecho de las grabaciones con las que hemos trabajado y de los datos obtenidos por otros investigadores proponemos para nuestro estudio el conjunto de rasgos faciales realizado por los hablantes para transmitir una información relacionada con el tipo de enunciado emitido: en la producción de enunciados declarativos encontraremos ausencia de fruncimiento del ceño, elevación de las cejas y ausencia de movimiento de la cabeza. Por su parte, durante la producción de los enunciados interrogativos conseguiremos fruncimiento del ceño, ausencia de elevación de las cejas y movimiento de la cabeza (a izquierda o derecha).

El conjunto de estos rasgos gestuales presentes en la producción de enunciados interrogativos y declarativos vendrían a constituir lo que Cavé (2003) llama *visemas* o lo que MacDonald y McGurk llaman *unidades de creencia audio-visuales*. Según Cavé (2003:63), *el efecto McGurk ha dado origen a una línea de investigación con miras a describir e identificar los «visemas», equivalentes visuales de los fonemas* (cf. figura 1).

1.1. Distinción acústica entre enunciados declarativos e interrogativos

Ahora bien, aunque en las distintas lenguas del mundo la distinción entre enunciados declarativos e interrogativos puede estar marcada por modificaciones morfosintácticas y de entonación, estar marcada morfosintácticamente o estar marcada solo por la entonación (Frota, Butler y Vigário, 2013), parece probable la idea de la existencia de una tendencia universal a usar una F0 alta para marcar los enunciados interrogativos y una F0 baja para marcar los enunciados declarativos (Soderstrom, Ko y Nevzorova, 2011; Yang y Liang, 2012). Gussenhoven y Chen (2000) indican que las personas tienden a percibir los enunciados con entonación alta y/o con tonos finales altos como interrogativos independientemente de la lengua materna. Sin embargo, Cangemi y D'Imperio (2013) señalan que las lenguas romance, germánicas y las no-indoeuropeas, además de mostrar contrastes claros entre los enunciados declarativos e interrogativos basados en las variaciones de la F0, marcan contrastes también a través de los cambios de la duración. Estas afirmaciones se comprueban en los señalamientos hechos por otros autores para el español: Astruc, Mora y Rew (2010:20) señalan que las interrogativas absolutas presentan sobre todo un tono más elevado que las declarativas. Méndez (2010:16) en su artículo con datos del proyecto AMPER refiere que *un menor campo tonal interactúa con un mayor campo durativo en la configuración de frases declarativas; en las interrogativas sucede lo contrario: menor campo durativo y mayor campo tonal*.



2. METODOLOGÍA

Partiendo del supuesto de que en la onda sonora de declarativas e interrogativas se encuentra codificada la distinción acústica entre estos dos tipos de enunciados, en este trabajo nos centraremos en el hecho de querer determinar hasta qué punto los oyentes hacen uso de la información visual para establecer una diferencia perceptiva entre esos dos tipos de enunciados. Creemos que el acceso a la información visual durante la percepción del habla influye sobre el resultado final del proceso de percepción.

2.1. Las grabaciones

Con la intención de determinar si la expresión facial influye en la percepción de la expresión hablada de los enunciados declarativos e interrogativos, se trabajó con una parte de un corpus para el estudio de la influencia de la información visual durante la percepción del habla recogido por integrantes y colaboradores del Laboratorio de Fonética de la Universidad de Los Andes. Específicamente trabajamos con un grupo constituido por 20 grabaciones audiovisuales: 10 de ellas contienen la emisión de enunciados declarativos y las otras 10 contienen los mismos enunciados pero emitidos de manera interrogativa. Así tenemos frases como *en Venezuela se habla español/ ¿En Venezuela se habla español? Su familia es de Mérida/ ¿su familia es de Mérida?*, entre otros. Los 20 enunciados fueron recogidos con 5 actores profesionales: dos hombres y tres mujeres con edades comprendidas entre los 20 y 24 años. A estas 20 grabaciones audiovisuales se les extrajo el audio con el fin de constituir un grupo de grabaciones de audio para el test auditivo.

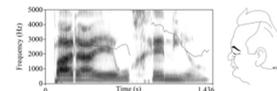
A continuación se presentan en la figura 1 dos imágenes de muestra extraídas de las grabaciones en las que se pueden apreciar los rasgos faciales distintivos producidos por uno de los actores mientras emitía la palabra *María* de la frase *María se vino desde Caracas solo para verlo* en sus dos modalidades: declarativa e interrogativa.



Figura 1. *Expresión facial de uno de los actores durante la emisión de la palabra María de la frase María se vino desde Caracas solo para verlo. A la izquierda la emisión declarativa y a la derecha la interrogativa.*

2.2. El test de percepción auditivo

Este test está constituido por 20 estímulos, correspondientes a los audios extraídos de las 20 grabaciones audiovisuales de enunciados declarativos e interrogativos. Para su



elaboración se distribuyeron aleatoriamente los 20 estímulos. Cada uno fue presentado dos veces, dejando 30 segundos de pausa entre ellos. Después de la segunda presentación, los participantes debían seleccionar de entre dos opciones posibles (declarativa o interrogativa) el tipo de enunciado percibido en cada uno de los estímulos. Este test se aplicó a varios grupos de personas de manera secuencial en el Laboratorio de Fonética de la Universidad de Los Andes. Para la aplicación se usaron un computador y unas cornetas a 65dB.

2.3. El test de percepción audiovisual

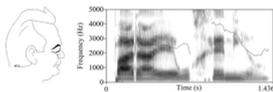
Una vez registradas, 10 de las 20 grabaciones audiovisuales fueron seleccionadas aleatoriamente y manipuladas con el programa Magic Video Deluxe: a cada una de ellas se les substituyó el audio original por el de otra. Así, si el enunciado registrado en la grabación era del tipo declarativo, se le substituyó el audio por el del mismo enunciado emitido de forma interrogativa, obteniendo con ello un estímulo con expresión facial de *declarativa* y audio de *interrogativa*. De esta manera, este test está constituido por 20 estímulos: 10 modificados y 10 no modificados. Al igual que en el test de percepción auditiva, a los participantes se le pidió que seleccionaran de entre dos opciones posibles (declarativa o interrogativa) el tipo de frase percibida en cada una de los estímulos. Las grabaciones fueron proyectadas a los participantes en una pantalla con un video beam y unas cornetas a 65dB. La aplicación de este test fue en las instalaciones del Laboratorio de Fonética de la Universidad de Los Andes.

2.4. Los participantes

Los tests de percepción fueron realizados por 200 personas (104 hombres y 96 mujeres), con un promedio de edad de 25,4 años, todos hablantes de español venezolano.

2.5. El procesamiento acústico

Cada una de las 20 grabaciones de audio fue procesada con el programa Praat para extraer los valores acústicos de frecuencia fundamental, intensidad y duración correspondientes a los enunciados declarativos e interrogativos con el fin de establecer los valores promedios de estos parámetros acústicos para cada tipo de enunciado. Pudimos apreciar que los enunciados declarativos presentan menor frecuencia fundamental (198,93 Hz) y mayor duración (2.04 seg.), mientras que lo contrario sucede con los enunciados interrogativos: presentan mayor frecuencia fundamental (210,21 Hz) y menor duración (1.81 seg.). Los valores de intensidad para ambos tipos de enunciados fueron muy similares (74,02 dB para los declarativos y 73,18 dB para los interrogativos). Estos datos vienen una vez más a confirmar lo que se ha sostenido en otras investigaciones para el español (cf. apartado 1.1.).



3. RESULTADOS Y DISCUSIÓN

3.1. El test de percepción auditivo

De los datos obtenidos para este test se desprende, en primer lugar, que todos los estímulos, independientemente de si son declarativos o interrogativos, fueron percibidos en altos porcentajes por los 200 participantes del test. El porcentaje de acierto es de 89,8. Los estímulos 1, 4 y 8 han sido los mejores percibidos con un 98% (es decir que 196 participantes del test percibieron correctamente estos estímulos declarativos), mientras que los estímulos 2, 12 y 5 han sido los que han obtenido el menor porcentaje de percepción correcta: 80% para los estímulos 2 y 12 y 78% para el estímulo 5 (esto equivale a decir que 160 participantes percibieron correctamente los estímulos 2 y 12 y solo 156 de los participantes han percibido correctamente el estímulo 5, todos contentivos de enunciados declarativos). Los enunciados declarativos han obtenido el mismo porcentaje de percepción correcta que los interrogativos (89,8%).

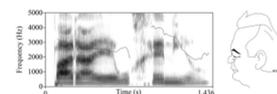
Con estos datos podemos afirmar que los participantes del test han podido percibir en la onda sonora de declarativos e interrogativos la distinción acústica entre estos dos tipos de enunciados, tal como lo señalan los datos acústicos comentados en el apartado 1.1. Los porcentajes de percepción acertada de los estímulos nos permiten decir que los participantes han utilizado los índices acústicos presentes en la señal. En consecuencia, tal como lo afirman Mora, Rojas, Méndez y Martínez (2008), la tarea perceptiva se realizó únicamente sobre la base de la señal acústica disponible.

3.2. El test de percepción audiovisual

Para la presentación de los resultados de este test discriminaremos entre los obtenidos para los estímulos no-modificados y los modificados, con el fin de determinar si hay alguna influencia de la información visual sobre la percepción de la información auditiva.

Luego de analizar los datos arrojados por el test, pudimos apreciar que los estímulos audiovisuales no-modificados son percibidos en altos porcentajes, al igual que sus correspondientes auditivos. Los estímulos 6 y 9 obtuvieron porcentajes altos de acierto al ser seleccionado como enunciado percibido el interrogativo (92 y 90% respectivamente). Por su parte, los estímulos audiovisuales 11, 16 y 14 obtuvieron los más bajos porcentajes de percepción acertada (75, 76 y 78% respectivamente). El hecho de que estos valores porcentuales obtenidos para los estímulos audiovisuales sean inferiores a los obtenidos para los mismos estímulos auditivos contradeciría el planteamiento formulado por Borrás, Kaland y Prieto (2013) y Crespo, Kaland, Swerts y Prieto (2013) según el cual la precisión de la percepción se mejora aún más cuando se añade información visual.

Por su parte, al apreciar los datos obtenidos para los estímulos modificados se puede apreciar que para el número 13 (input visual correspondiente al enunciado emitido de manera interrogativa e input sonoro correspondiente al mismo enunciado emitido de manera declarativa), 188 participantes (94%) escogieron como opción el enunciado interrogativo, y 12 participantes (6%), el tipo declarativo como enunciado percibido.



Datos similares se han obtenido para el estímulo 15 (input visual interrogativo e input sonoro declarativo): 186 participantes (93%) escogieron el tipo interrogativo y 14 (7%) el tipo declarativo; y para el estímulo 7 (input visual interrogativo e input sonoro declarativo): 184 participantes (92%) escogieron el tipo interrogativo y 16 (8%) el tipo declarativo. Al analizar cada uno de estos casos, vemos que los participantes del test han basado su decisión perceptiva en el input visual y no en el auditivo. Estos rasgos nos permiten afirmar que la configuración facial de las frases interrogativas (fruncimiento del ceño e inclinación de la cabeza y la no elevación de cejas) tiene un peso importante a la hora de tomar estas decisiones perceptivas.

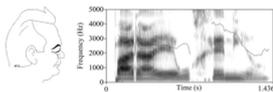
Esta preferencia por el input visual también puede apreciarse en los datos obtenidos para los estímulos 18, 1 y 5. En el primero de ellos (input visual declarativo e input sonoro interrogativo) pudimos apreciar que la tendencia es a percibirlo como declarativo, 180 (90%) participantes así lo han decidido, mientras que 20 (10%) participantes lo han percibido como interrogativo. Los mismos datos se han obtenido para el estímulo 1 (input visual declarativo e input sonoro interrogativo). 164 (82%) de los participantes de nuestro test han percibido que el estímulo 5 (input visual declarativo e input sonoro interrogativo) también contiene un enunciado declarativo. Los rasgos presentes en la configuración facial que aparece durante la emisión de enunciados declarativos (la elevación de cejas y ausencia de fruncimiento del ceño y de la inclinación de la cabeza) tiene un valor fundamental para los perceptores.

Por su parte, en el estímulo 4 de nuestro test (input visual declarativo e input sonoro interrogativo), 178 participantes (89%) escogieron el tipo interrogativo como enunciado percibido en el estímulo, mientras que 22 personas (11%) escogieron como enunciado percibido una declarativo. Este mismo hecho ocurrió con dos estímulos: el estímulo 3 (input visual declarativo e input sonoro interrogativo), en el que 176 (88%) participantes percibieron el estímulo como interrogativo y 24 (12%) lo percibieron como declarativo; y el estímulo 17 (input visual declarativo e input sonoro interrogativo): 182 (91%) personas lo percibieron como interrogativo mientras que 18 (9%) lo percibió como declarativo. De estos datos se desprende que los participantes basaron la percepción de estos estímulos en lo sonoro y no en lo visual. A pesar de que en los videos de cada uno de estos estímulos la configuración facial se correspondía con la de la emisión de un enunciado declarativo (caracterizados por la elevación de cejas, ausencia de fruncimiento del ceño y de la inclinación de la cabeza), la información sonora tuvo mayor peso a la hora de tomar una decisión perceptiva.

De acuerdo con los resultados de nuestro test, los perceptores atienden en primer lugar a la información visual por sobre la información acústica. Cuando la información visual producida durante la emisión del enunciado no es clara para los perceptores, atienden entonces a la información acústica, tal como lo señalan Martínez, Rojas y Suárez (2012).

4. CONCLUSIONES

Nuestro objetivo inicial era determinar hasta qué punto los oyentes hacen uso de la información visual para establecer la categoría gramatical a la que pertenecen los



enunciados declarativos e interrogativos. Luego de la aplicación de los tests afirmamos junto con otros investigadores (cf. Soto, 2000; Borrás, Kaland y Prieto, 2013; Crespo, Kaland, Swerts y Prieto, 2013, por ejemplo) que el acceso a la información visual durante la percepción del habla influye sobre el resultado final del proceso de percepción del habla.

Hemos dicho que en la onda sonora de declarativas e interrogativas se encuentra codificada la distinción acústica entre estos dos tipos de enunciados. Sin embargo, cuando los perceptores cuentan con información visual asociada a esa onda sonora de declarativas e interrogativas, esa información puede modificar el resultado del proceso perceptivo contundentemente. Así, a través de nuestros datos hemos logrado establecer una relación entre el input visual y el auditivo durante la percepción de este tipo de enunciados.

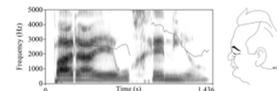
Estos hechos nos llevan a sugerir junto con Collignon et al. (2008), Robinson y Sloutsky (2004), Swerts y Kraemer (2005) y Dijkstra et al. (2006) que la percepción visual es indudablemente dominante, pero debemos sugerir también que este dominio no ocurre de manera rígida, sino flexible. Esta afirmación nos lleva a agregar, junto con Martínez, Rojas y Suárez (2012), que el producto del proceso de percepción audiovisual es una correlación del procesamiento de las informaciones visual y sonora. Los perceptores están cotejando siempre ambos inputs durante el proceso de percepción del habla. Por tanto, confirmamos que la atención entonces recaerá en aquel input que mejor se adapte a las expectativas (contextuales/situacionales, culturales, sociales, anímicas, etc.) del perceptor. Siguiendo a Massaro (1987:85), tanto la información auditiva como la visual influyen en la percepción de la prosodia: *la influencia de una modalidad (auditiva o visual) será mayor cuando la información proporcionada por la otra modalidad sea ambigua.*

AGRADECIMIENTOS

Esta investigación fue realizada con el financiamiento del Consejo de Desarrollo Científico, Humanístico, Tecnológico y de las Artes (CDCHTA) de la Universidad de Los Andes a través del proyecto H-1438-13-06-B.

5. REFERENCIAS BIBLIOGRÁFICAS

- ASTRUC, LL.; E. MORA y S. REW. (2010): «Venezuelan Andean Spanish Intonation», en P. Prieto y P. Roseano (eds.): *Transcription of Intonation of the Spanish Language*, Munich, Lincom Europa, pp. 191-226.
- BORRÁS, J.; C. KALAND; P. PRIETO y J. SWERTS (2013): «Audiovisual correlates of interrogativity: A comparative analysis of Catalan and Dutch», *Journal Nonverbal Behavior*, pp. 125-142.
- CAVE, CH. (2003): «Illusions auditives», *Dossier Pour la Science*, 39, pp. 60-63.
- CANGEMI, F. y M. D'IMPERIO (2013): «Tempo and perception of sentence modality», *Journal of the Association for Laboratory Phonology*, 4, 1, pp. 191-219.
- COLLIGNON, O.; S. GIRARD; F. GOSSELIN; S. ROY; D. SAINT-AMOUR; M. LASSONDE y F. LÉPORE. (2008): «Audio-visual integration of emotion expression», *Brain research*, 1242, pp. 126-135.



- CRESPO, V.; C. KALAND; M. SWERTS y P. PRIETO. (2013): «Perceiving incredulity: The role of intonation and facial gestures», *Journal of pragmatics*, 47, pp. 1-13.
- DIJKSTRA, C.; E. KRAHMER y M. SWERTS. (2006): «Manipulating uncertainty: The contribution of different audiovisual prosodic cues to the perception of confidence», en R. Hoffmann y H. Mixdorff (eds.): *Proceedings of the 3rd International Conference on Speech Prosody*, Desden (Alemania).
<http://arno.uvt.nl/show.cgi?fid=95687> [12/03/2012].
- FROTA, S.; J. BUTLER y M. VIGÁRIO. (2013): «Infants' perception of intonation: Is it a statement or a question?», *Infancy*, 19, 2, pp. 194-213.
<http://dx.doi.org/10.1111/infa.12037> [12/03/2012].
- GUSSENHOVEN, C. y A. CHEN. (2000): «Universal and language-specific effects in the perception of question intonation», en J. Ohala, Y. Hasegawa, M. Ohala, D. Granville y A. Bailey (eds.): *International Conference on the Processing of Spoken Language*, 6, 2, pp.1-4.
- HIGUERA, L. (2013): *Influencia de la información visual durante la percepción de frases interrogativas y declarativas del español venezolano*, memoria de gradom Universidad de Los Andesm Mérida.
- MARTÍNEZ, H.; D. ROJAS y F. SUÁREZ (2012): «Influencia de la información visual durante la percepción de la prosodia de las emociones actuadas», *Estudios de Fonética Experimental*, XXI, pp.163-193.
- MASSARO, D. (1987): *Speech perception by ear and by eye*, Hillsdale, N J, Erlbaum.
- MÉNDEZ, J. (2010): «Interacción de los parámetros acústicos duración y frecuencia fundamental en frases declarativas neutras e interrogativas absolutas de los andes venezolanos», *Estudios de Fonética experimental*, XIX, pp.147-164.
- MORA, E.; N. ROJAS; J. MÉNDEZ y H. MARTÍNEZ (2008): «Declarativas e interrogativas del español venezolano. Percepción de la emisión con y sin contenido léxico», *Language Design*, 2, pp. 231-238.
- ROBINSON, C. y V. SLOUTSKY (2004): «Auditory dominance and its change in the course of development», *Child Development*, 75, 5, pp. 1387–1401.
- SODERSTROM, M.; E. KO y U. NEVZOROVA. (2011): «It's a question? Infants attend differently to yes/no questions and declaratives», *Infant Behavior and Development*, 34, pp. 107-110.
<http://dx.doi.org/10.1016/j.infbeh.2010.10.003> [12/03/2012].
- SOTO, M. (2000): *Influencia de la percepción visual del rostro del hablante en la credibilidad de su voz*, tesis doctoral, Universitat Autònoma de Barcelona.
- SWERTS, M. y E. KRAHMER. (2005): «Audiovisual prosody and feeling of knowing», *Journal of Memory and Language*, 53(1), pp. 81–94.
- YANG, X. y J. LIANG. (2012): «Declarative and interrogative intonations by brain-damaged speakers of Uygur and Mandarin Chinese», en Q. Ma, H. Ding y D. Hirst (eds.): *Proceedings of the Sixth International Conference on Speech Prosody*, Shanghai, China, pp. 701-704.