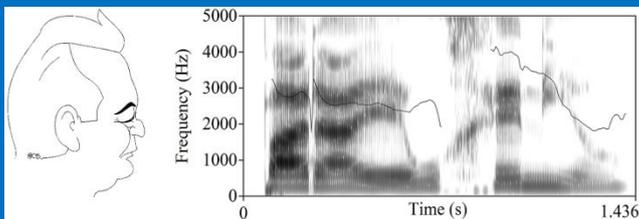


Fonética articulatorio-auditiva: evidencia (localización) cerebral del modelo de integración sensorio-motora

Faustino Diéguez Vide



Fernández Planas, A. Ma. (ed.) (2016): *53 reflexiones sobre aspectos de la fonética y otros temas de lingüística*, Barcelona, págs. 343-352.

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

Fonética articulatorio-auditiva: evidencia (localización) cerebral del modelo de integración sensorio-motora

Faustino Diéguez Vide
Universitat de Barcelona
fdiequez@ub.edu

Aunque se hable de tres fonéticas, no hay que pensar que realmente son tres cosas completamente separadas. [...]Esas tres partes son solo tres aspectos o puntos de vista del mismo fenómeno. En cada momento se podrá adoptar el punto de vista más adecuado para describir un fenómeno, pero forzosamente esa perspectiva estará relacionada con los otros dos aspectos.

E. Martínez Celdrán, *Fonética* (19847), Barcelona, Teide.

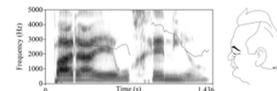
1. INTRODUCCIÓN

La década de los noventa, el siglo XX y la primera década del siglo XXI han permitido desmentir –o, al menos, poner en duda– diferentes presupuestos teóricos y empíricos en muchos ámbitos diversos. La Lingüística es uno de estos campos y, dentro de esta, la Fonética: así, por ejemplo, la noción teórica de «fonema» se ha puesto en evidencia al no existir una realidad psicológica del mismo, o, por lo menos, una realidad psicológica del mismo en las lenguas románicas. La Neurociencia es otro de los campos y, dentro de esta, la Neurolingüística: por ejemplo, la modularidad fodoriana, presente en muchos trabajos de la década de los noventa, no se mantiene hoy en día empíricamente.

En este escrito se intentará poner en duda un nexo de unión entre los ámbitos comentados anteriormente a través de una revisión de la literatura. Los estudiantes de Lingüística pronto se dan cuenta de que lo que se les enseña dista mucho de ser lo «realmente cierto»: la producción y la comprensión del lenguaje van mucho más allá de Broca y de Wernicke, nociones básicas y axiomas sobre el lenguaje se desmoronan empíricamente (¿existe el fonema?, ¿la sintaxis es jerárquica o lineal?, etc.). No obstante, en las disciplinas científicas –al menos en nuestro ámbito– no suele producirse un estudio de los vínculos entre disciplinas: se estudia académicamente la Morfología y la Sintaxis, pero qué pasa con la Morfosintaxis. En las líneas que siguen, se intentará proponer un vínculo entre Fonética y Neurolingüística: se relacionarán las tareas propias de la fonética articulatoria y de la fonética auditiva pues, siempre, actúan de forma coordinada, al menos en el cerebro.

2. LOS INICIOS (PLANTEAMIENTO DEL TEMA)

Todos los modelos «clásicos» (es decir, desde sus inicios en la década de los sesenta hasta los modelos de corte mental más recientes) relacionados con la comprensión y la



producción orales siempre han incluido un componente periférico para tratar ambos aspectos. En concreto, se han propuesto tres posibilidades:¹

1. Los denominados modelos de dos rutas plantean una entrada periférica para la audición (denominada «análisis acústico auditivo» o «sistema de análisis auditivo») y una salida periférica para la articulación (que en este caso siempre se suele denominar «articulación»). Los recientes descubrimientos de la tractografía han revitalizado estos modelos (Coltheart et al, 2001).
2. Muchos modelos de ámbito mental plantean también la misma entrada y la misma salida periféricas, pero sin suponer la existencia de dos rutas separadas. Serían modelos como el de Levelt (1983) para la descripción de la producción y la comprensión orales, Garrett (1980) para la producción, y Forster (1976) o Morton (1979) para la comprensión.
3. Los modelos de carácter interactivo o conexionista igualmente plantean una entrada auditiva para la comprensión oral y una salida articular para la producción oral, pero se incluye más de un procesamiento. Así, por ejemplo, el modelo TRACE de McClelland y Elman (1986) para la comprensión incluye, además de un nivel de fonemas, otro nivel anterior de rasgos fonéticos; por su parte, el modelo de Stemberger (1985) para la producción oral conecta de forma interactiva un nivel fonológico con un nivel de rasgos fonéticos. El interés de estos modelos es que incorporan no solo conexiones de activación entre niveles, sino también conexiones de inhibición, así como el denominado procesamiento en cascada y la noción de decaimiento, es decir, del retorno de una conexión a su estado de reposo inicial (Dell, 1986; Martin et al, 1994).

A pesar de la evidencia empírica ofrecida por todos estos modelos, en forma de experimentos controlados de laboratorio o a partir de los *lapsus linguae*, los recientes avances en neuroimagen han mostrado que el procesamiento cerebral es algo más complejo, aunque algunos modelos conexionistas poseen características similares a las de esta propuesta. Como apuntan Hickok, Houde y Rong (2011), los cambios introducidos en los modelos actuales, a partir de la evidencia cerebral, son dos: primero, que el sistema auditivo está críticamente implicado en la producción del habla; y, segundo, que el sistema articular o motor está críticamente implicado en la percepción del habla. Dada la complejidad de ambos temas, aquí solo se trabajará el ámbito de la producción oral.

Es bastante fácil darse cuenta de que nuestro sistema auditivo nos influye en la producción del habla. Más allá de diversos experimentos psicolingüísticos que así lo demostrarían, es posible observar evidencias en diversas actividades diarias: somos capaces de autocorregirnos cuando al hablar cometemos algún error, necesitamos que el canal de la comunicación esté siempre ocupado con una recepción (como al hablar por teléfono), los sujetos que adquieren una sordera sufren un declive articular y los sujetos con patología congénita no articulan igual, los *lapsus linguae* no se producen con

¹ Muchos de los modelos presentados aquí están compilados en castellano en Valle, Cuetos, Igoa y del Viso (1990), por lo que se remite al lector a esta obra. En concreto: Forster (1976), Morton (1979), una versión del TRACE (McClelland y Elman, 1986), Garrett (1980) y Stemberger (1985).



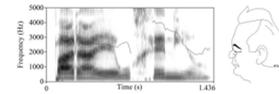
desajustes fonéticos... ¿por qué no hablamos una segunda lengua en voz alta cuando estamos aprendiendo esa lengua?

Existe un punto de partida claro en relación con la vinculación comentada: los trabajos de Hickok y colaboradores (Hickok et al, 2003; Hickok, Okada y Serences, 2009; Pa y Hickok, 2008; Buchsbaum et al, 2011) en los que se ha propuesto una región anatómica que permitiría la integración auditivo-motora cuando se articula una emisión: la parte posterior del *planum temporale* (PT), localizado debajo del área de Wernicke y su extensión hacia el opérculo parietal (OP) en las inmediaciones subcorticales del área de Broca. Esta área, denominada Spt (de localización silviana entre los lóbulos parietal y temporal), estaría lateralizada hacia el hemisferio izquierdo, y estaría activa incluso en ausencia de actividad motora o sensorial: por ejemplo, en conducta exclusiva de percepción o cuando se realiza una tarea cerrada (*covert*), es decir, en silencio. Lo interesante es que esta área recibe dos tipos de influencias, una de pre-alimentación (o *feedforward*) por parte principalmente de estructuras frontales –es decir, motoras o articulatorias–, y otra de retroalimentación (o *feedback*) por parte de áreas temporales y parietales relacionadas, respectivamente, con el procesamiento auditivo y somato-sensorial. La relación con estructuras motoras es especialmente significativa con la corteza premotora, ambos cerca del área de Broca (Hickok et al, 2003; Buschmaun et al, 2011). En relación con las áreas posteriores, es también relevante la relación con la circunvolución supramarginal, en el lóbulo parietal (Price, 2010), y con la circunvolución temporal superior (donde se encuentra el área de Wernicke). Esta compleja red se completaría con la participación de la ínsula, el cerebelo y áreas subcorticales. Es necesario advertir el hecho de que esta área no es específica del lenguaje, pues también responde ante estímulos tonales o melódicos.

3. EL MODELO DE INTEGRACIÓN SENSORIO-MOTOR EN MONOLINGÜES

Tras el trabajo comentado en el párrafo anterior, surgieron diferentes propuestas sobre la influencia de la audición en la producción del habla y comenzaron también a observarse algunas divergencias. Estas estarían originadas por diversos factores, como apunta el grupo de Simmonds (Simmonds et al, 2014a, 2014b). Así, el uso de paradigmas con habla abierta (*overt*) o cerrada (*covert*), es decir, con producción oral en el contexto de lectura en voz alta o con lectura silenciosa, ha llevado a algunas conclusiones diferentes (Yetkin et al, 1995; Barch et al, 1999; Palmer et al, 2001). De igual forma, podrían existir resultados divergentes por problemas metodológicos, como una muestra pequeña de participantes o una falta de continuidad en la imagen. En este último caso, se han utilizado tareas tan dispares como la prueba Stroop (Barch et al, 1999), el completamiento de palabras (Palmer et al, 2001) o tareas de fluencia verbal (Yetkin et al, 1995).

Pero más allá de estas divergencias, sí existen algunos datos comunes. Las dos estructuras centrales en el procesamiento de la articulación serían, como ya se ha comentado, el *planum temporale* (PT) y el opérculo parietal. La relación de estas áreas con los mecanismos de control *feedforward* y *feedback* puede apreciarse en la figura de Golfinopoulous, Tourville y Guenthe (2010; ver página siguiente) que, a pesar de reflejar el funcionamiento de su denominado modelo DIVA (*Directions Into Velocities of*



Articulators), muestra claramente las áreas implicadas (las abreviaturas se describen en el texto, con las iniciales inglesas).

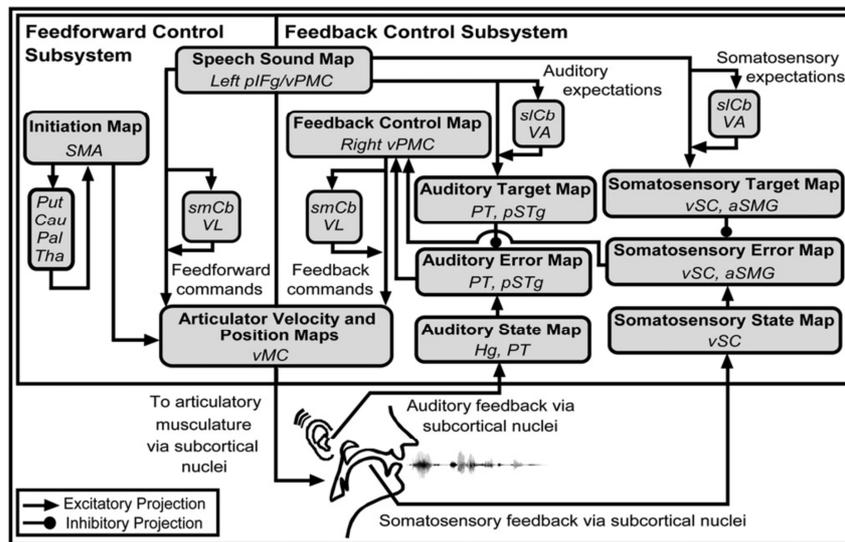
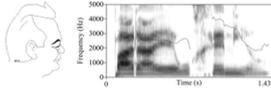


Figura 1. Modelo DIVA de la producción articulatoria.

Así, en relación con las áreas de control *feedforward*, es decir, las áreas motoras, existiría una red formada por estructuras frontales, subcorticales y cerebelosas. Existen dos estructuras que funcionarían como un puente vinculando las áreas motoras con las áreas auditivas y sensoriales: el complejo formado por la parte ventral de la corteza premotora (vPMC), la parte posterior de la corteza frontal inferior izquierda (que incluye el área de Broca: pIFg) y la parte ventral de la corteza motora (vMC). Participarían también en la producción el área motora suplementaria (SMA) que se encargaría de la planificación de la emisión, y un conjunto de estructuras subcorticales [los ganglios basales –núcleo caudado (Cau), globo pálido (Pal) y putamen (Put)– y el tálamo (Tha)] y cerebelosas (núcleo ventral lateral de la superficie superomedial del cerebelo: smCB VL).

Por su parte, las estructuras de *feedback* serían de dos clases, auditivas y somato-sensoriales. Existiría un área de control del *feedback*: la parte ventral de la corteza premotora derecha. Igual que para el *feedforward*, en el *feedback* participan estructuras cerebelosas, tanto en la audición como en la sensación: en concreto, el núcleo ventral lateral de la superficie superomedial del cerebelo y el núcleo anterior de la superficie superolateral del cerebelo (slCB VA). Desde una perspectiva más específica en el *feedback* auditivo participan el PT, la circunvolución de Heschl (HG) y la parte posterior de la circunvolución temporal superior o primera (pSTG), a lo que debería añadirse el surco temporal superior. Por su parte, en el *feedback* sensorial participan la parte ventral de la corteza somatosensorial (vCS) y la parte anterior de la circunvolución supramarginal (aSMG).

Así, toda esta red se pone en funcionamiento cuando las personas producimos emisiones orales y se recoge en esta propuesta la relación existente entre dos sistemas fonológicos,



uno motor y otro auditivo. La activación de la red al producir una emisión implica una activación simultánea en los sistemas fonológico motor y auditivo. La activación del componente auditivo involucra los objetivos sensoriales de la acción, mientras que la activación del sistema fonológico motor define el plan motor inicial, mediante *feedbacks* internos que permitirán una comparación con aquellos objetivos. Toda la zona temporal comentada (circunvolución temporal superior y surco temporal superior) se correspondería con el procesamiento fonológico auditivo, mientras que las zonas frontales (corteza pre-motora) se relacionarían con el procesamiento fonológico motor; la relación entre estos dos sistemas fonológicos estaría mediada por el área Spt.

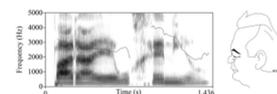
No obstante, existen algunas investigaciones que, trabajando con paradigmas que utilizan tareas abiertas, han mostrado una localización algo diferente de las áreas relacionadas con el modelo de integración sensorio-motor. El grupo de Simmonds, Wise, Leech y colaboradores (Simmonds et al, 2014a, 2014b) han mostrado que en tareas abiertas, el procesamiento prearticulatorio de la producción en el lóbulo frontal estaría fuertemente lateralizado en el hemisferio izquierdo (en el área de Broca y de acuerdo con los estudios previos), pero el flujo de la corteza motora primaria que controla los músculos axiales y, sobre todo, el *feedback* auditivo y sensorial estarían localizados en ambos hemisferios. Así, por ejemplo, en una tarea abierta de generación de oraciones, Tremblay y Small (2011) han observado una activación bilateral en las circunvoluciones transversas de Heschl, con una extensión caudal hacia el PT. Además, en estos estudios se ha observado también activación de otras áreas cerebrales: la zona anterior del área motora suplementaria y su extensión hacia la corteza cingulada anterior y la parte posterior de la circunvolución temporal media, junto con la parte anterior del PT (y no todo el PT).

4. EL MODELO DE INTEGRACIÓN SENSORIO-MOTOR EN BILINGÜES

Cuando se estudia el acento de hablantes bilingües en la lengua no-materna o de menor uso (L2), se observa un habla no-nativa (Díaz et al, 2008; Simmonds et al, 2011; Goldrick, Runnqvist y Costa, 2014). Se ha propuesto que solo el 15% de las personas adultas que aprenden una L2 son capaces de adquirir el acento de esa lengua hasta el punto de considerarse nativos. Esta actuación depende de diversas variables (ver Piske, MacKay y Flege, 2001, para una revisión), pero se ha observado que el habla no-nativa es especialmente pronunciada en sujetos con una adquisición tardía de la L2 (Sebastian-Gallés et al, 2006; Simmonds et al, 2011), aunque puede también observarse en sujetos que han tenido contacto con ambas lenguas desde la infancia (Scovel, 1988; Sebastian-Gallés, Echeverría y Bosch, 2005).

La aplicación del modelo de integración sensorio-motor en sujetos bilingües ha provocado que algunos de los supuestos comentados en el apartado anterior se hayan visto comprometidos. Estas divergencias se pueden integrar en dos aspectos: uno funcional y otro relacionado con las variables comentadas.

1. En primer lugar, hay dos trabajos (Simmonds, Wise y Leech, 2011; Simmonds et al, 2001) en los que se afirma un funcionamiento diferenciado del modelo para la L1 y para la L2. Se aducen dos razones:

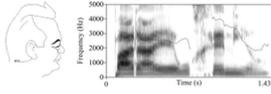


- a. Los articuladores que controlan el habla utilizan los mismos músculos que usamos para respirar: si esos controles dependen del *feedback* auditivo, los hablantes de una L2 necesitan una destreza especial para controlar el movimiento motor necesario para respirar, para articular la L1 y para articular la nueva L2. Esta destreza se logra, normalmente, en aquellos casos en que un sujeto no solo adquiere o aprende una L2 sino que se traslada a un nuevo lugar y solo habla esa L2.
 - b. La producción de una L1 es altamente automática, lo que implica una producción oral sin apenas errores y un mínimo control sensorial. El hecho de que la neuroimagen muestre una señal reducida o ausente en la corteza sensorial mientras se emite habla en la L1 indicaría un incremento en la eficiencia del *feedback* gracias a la sincronización del sistema lograda por la experiencia del habla desde la infancia. Por el contrario, la producción en una L2 es menos automática y se observan muchos más errores, lo que implica una participación activa del control sensorial y auditivo vía *feedback*, un control que debe eliminar las discrepancias entre el resultado sensorial de los movimientos pretendidos y reales.
2. Pero esto no sucede siempre. El estudio comentado ha mostrado un incremento de la actividad en la L2 solo en bilingües tardíos, es decir, en sujetos que han aprendido la L2 después del denominado período crítico (que se suele situar, dependiendo de los autores, entre los 5 y los 7 años). Y este incremento aparece, incluso, en sujetos que, habiendo adquirido tardíamente la L2 poseen una alta competencia en esa L2. A pesar de la alta competencia, la L2 tendrá «acento».

Desde estos postulados, entonces, se admite una actuación diferenciada entre monolingües, por un lado, y bilingües, por otro. Pero, además, también una actuación diferenciada entre hablantes bilingües tempranos y tardíos, con independencia de la competencia adquirida.

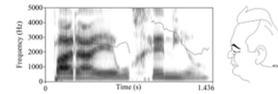
5. CONCLUSIONES

Iniciábamos esta andadura comentando la importancia de las vinculaciones entre disciplinas y abogábamos por una implicación de algunos aspectos fonéticos con la localización cerebral. En este trabajo se ha presentado un modelo sobre la articulación en la producción oral y se ha mostrado que esta conducta requiere de una compleja red cerebral en la que participan estructuras de control *feedforward* y *feedback*. Igualmente, se han mencionado las posibles estructuras cerebrales que participan en esa compleja red: se parte de un sistema de procesamiento fonológico motor (articulatorio) y de otro sistema de procesamiento fonológico auditivo (percepción), unidos por un sistema de traducción. Desde aquí, se proponen tres propuestas para el futuro: primero, cómo describe este modelo diversas patologías; segundo, cómo influyen las diversas variables relacionadas con el bilingüismo en el procesamiento de estos sistemas; y, tercero, cómo ampliar el modelo para integrar también conductas lingüísticas más allá de la articulación.



6. REFERENCIAS BIBLIOGRÁFICAS

- BARCH, D.M.; F. W. SABB; C. S. CARTER; T. S. BRAVER; D. C. NOLL y J. D. COHEN (1999): «Overt verbal responding during fMRI scanning empirical investigations of problems and potential solutions», *Neuroimage*, 10, pp. 642-657.
- BUCHSBAUM, B.R.; J. BALDO; K. OKADA; K. F. BERMAN; N. DRUNKERS; M. D'ESPOSITO y G. HICKOK (2011): «Conduction aphasia, sensory-motor integration, and phonological short-term memory – an aggregate analysis of lesion and fMRI data», *Brain and Language*, 119, pp. 119-128.
- COLTHEART, M.; K. RASTLE; C. PERRY; R. LANGDOM y J. ZIEGLER (2001): «DRA: A dual-route cascaded model of visual word recognition and reading aloud», *Psychological Review*, 108, pp. 204-256.
- VALLE, F.; F. CUETOS; J. M. IGOA y S. VISO (comp) (1990): *Lecturas de psicolingüística, vol. 1: Comprensión y producción del lenguaje*, Madrid, Alianza.
- DELL, G. (1986): «A spreading-activation theory of retrieval in sentence production», *Psychological Review*, 93, pp. 283-321.
- DIAZ, B.; C. BAUS; C. ESCERA; A. COSTA y N. SEBASTIÁN-GALLÉS (2008): «Brain potentials to native phoneme discrimination reveal the origin of individual differences in learning the sounds of a second language», *PNAS, Proceedings of the National Academy of Sciences*, 105, pp. 16083-18088.
- GOLDRICK, M.; E. RUNNQVIST y A. COSTA (2014): «Language switching makes pronunciation less nativelike», *Psychological Science*, 25, pp. 1031-1036.
- GOLFINOPOULOUS, E.; J. A. TOURVILLE y F. H. GUENTHE (2010): «The integration of large-scale neural network modeling and functional brain imaging in speech motor control», *Neuroimage*, 52, pp. 862-874.
- HICKOK, G.; K. OKADA y J. T. SERENCES (2009): «Area Spt in the human planum temporale supports sensory-motor integration for speech processing», *Journal of Neurophysiology*, 191, pp. 2725-2732.
- HICKOK, G.; J. HOUDE y F. RONG (2011): «Sensorimotor integration in speech processing: Computational basis and neural organization», *Neuron*, 69, pp. 407-422.
- HICKOK, G.; B. BUCHSBAUM; C. HUMPHRIES y T. MUFTULER (2003): «Auditory-motor interaction revealed by fMRI: Speech, music and working memory in area Spt», *Journal of Cognitive Neuroscience*, 15, pp. 673-682.
- LEVELT, W. J. (1983): «Monitoring and self-repair in speech», *Cognition*, 14, pp. 41-104.
- MARTIN, N.; G. S. DELL; E. M. SAFFRAN y M. F. SCHWARTZ (1994): «Origins of paraphasias in deep dysphasia: Testing the consequences of a decay impairment to an interactive spreading activation model of language», *Brain and Language*, 47, pp. 609-660.
- PA, J. y G. HICKOK (2008): «A parietal-temporal sensory-motor integration area for the human vocal tract: Evidence from a fMRI study of skilled musicians», *Neuropsychologia*, 46, pp. 362-368.
- PALMER, E. D.; H. J. ROSEN; J. G. OJEMANN; R. L. BUCKNER; W. M. KELLEY y S. E. PETERSEN (2001): «An event-related fMRI study of overt and covert word stem completion», *Neuroimage*, 14, pp. 182-193.
- PISKE, T.; I. R. A. MACKAY y J. E. FLEGE (2001): «Factors affecting degree of foreign accent in an L2: A review», *Journal of Phonetics*, 29, pp. 191-215.
- PRICE, C. J. (2010): «The anatomy of language: A review of 100 fMRI studies published in 2009», *Annals of the New York Academy of Sciences*, 1191, pp. 62-88.
- SCOVEL, T. (1988): «A time to speak: A psycholinguistic inquiry into the critical period for human speech», New York, Newbury House Publishers.
- SEBASTIÁN-GALLÉS, N.; S. ECHEVERRÍA y L. BOSCH (2005): «The influence of initial exposure on lexical representation: Comparing early and simultaneous bilinguals», *Journal of Memory and Language*, 52, pp. 240-255.
- SEBASTIÁN-GALLÉS, N.; A. RODRÍGUEZ-FORNELLS; R. DE DIEGO BALAGUER y B. DÍAZ (2006): «First- and second-language phonological representations in the mental lexicon». *Journal of Cognitive Neuroscience*, 18, pp. 1277-1291.



- SIMMONDS, A. J.; R. J. S. WISE y R. LEECH (2011): «Two tongues, one brain: Imaging bilingual speech production», *Frontiers in Psychology*, 2, 166.
- SIMMONDS, A. J.; R. J. S. WISE; N. S. DHANJAL y R. LEECH (2011): «A comparison of sensory-motor activity during speech in first and second languages», *Journal of Neurophysiology*, 106, pp. 470-478.
- SIMMONDS, A. J.; R. LEECH; C. COLLINS; O. REDJEP y R. J. S. WISE (2014a): «Sensory-motor integration during speech production localizes to both left and right plana temporale», *Journal of Neuroscience*, 24, pp. 12963-12972.
- SIMMONDS, A. J.; R. J. S. WISE; C. COLLINS; O. REDJEP; D. SHARP; P. IVERSON y R. LEECH (2014b): «Parallel systems in the control of speech», *Human Brain Mapping*, 35, pp. 1930-1943.
- TREMBLAY, P. y S. L. SMALL (2011): «Motor response selection in overt sentence production: A functional fMRI study», *Frontiers of Psychology*, 2, 253.
- YETKIN, F. Z.; T. A. HAMMEKE; S. J. SWANSON; G. L. MORRIS; W. M. MUELLER; T. I. MCAULIFFE y V. M. HAUGHTON (1995): «A comparison of functional MR activation patterns during silent and audible language tasks», *American Journal of Neuroradiology*, 16, pp. 1087-1092.