



Commentary: Visual Feedback of Tongue Movement for Novel Speech Sound Learning

Marcelo L. Berthier^{1*} and Ignacio Moreno-Torres²

¹ Cognitive Neurology and Aphasia Unit and Cathedra ARPA of Aphasia, Centro de Investigaciones Médico-Sanitarias, Instituto de Investigación Biomédica de Málaga, University of Malaga, Malaga, Spain, ² Department of Spanish Language, University of Malaga, Malaga, Spain

Keywords: visual feedback, auditory feedback, tongue movements, speech learning, neuroimaging, cholinergic systems

A commentary on

Visual Feedback of Tongue Movement for Novel Speech Sound Learning

by Katz, W. F., and Mehta, S. (2015). *Front. Hum. Neurosci.* 9:612. doi: 10.3389/fnhum.2015.00612

Access to visual information during speech is useful for verbal communication because it enhances auditory perception, especially in noise (Sumbly and Pollack, 1954). Until recently, visual feedback of the speaker's face was possible only for the visible articulators (the lips), but new computer-assisted pronunciation training (CAPT) systems have created a kind of "transparent" oral cavity that permits us to look at our own tongue movements during articulation.

The use of CAPT systems is a clever way to integrate visual feedback with auditory feedback. This has been exploited by Katz and Mehta (2015) (hereafter K & M) to strengthen learning of novel speech sounds. Using this emerging methodology (*Opti-Speech* visual feedback system—Katz et al., 2014), they trained five college-age subjects to learn a novel consonant in the /aCa/ context. Before training, the stimulus was produced by one of the investigators (SM) three times together with an explanation on how to position the tongue inside the mouth while reproducing the phoneme. The learned consonant was a sound not attested as a phoneme among the world's languages. The stimuli were elicited in blocks of 10 /aCa/ productions using a single-case ABA design (pre-training, training, post-training), with visual feedback only in the training phase. In their study, K & M presented the listeners with a visual representation of internal (hidden to vision) articulators to enhance perception and eventually to finely tune articulatory tongue movements. Five sensors were positioned on the tongue of the learners so as to provide them with online visual feedback of the hidden articulators thus allowing tongue reading. Despite receiving detailed verbal instructions, all subjects did poorly at baseline assessment; their accuracy improved during the visual feedback training and gains were maintained in a follow-up examination (both $p < 0.001$ relative to baseline). The effect size of training for each subject was $>90\%$ (highly effective). Analysis of acoustic and spectral parameters suggested increased production consistency after training.

Nevertheless, K & M's findings should be interpreted with caution as the study had some methodological drawbacks, which limit firm conclusions on the reported effectiveness of the training procedure. These include the small sample size, the uncontrolled nature of the study, the fact that not all participants completed the initial protocol, the use of a novel consonant (not attested as a phoneme in any natural language), and the unreliable replicability of target phoneme produced by the experimenter. These factors need to be controlled in future studies to ascertain whether motor learning of tongue movements with the support of 3D information is useful to learn speech sound in real life situations (L2 learning and speech-language rehabilitation). In any case,

OPEN ACCESS

Edited by:

Peter Sörös,
University of Oldenburg, Germany

Reviewed by:

Ana Inès Ansaldo,
Université de Montréal, Canada
Sandro Franceschini,
Università di Padova, Italy

*Correspondence:

Marcelo L. Berthier
marberthier@yahoo.es

Received: 19 September 2016

Accepted: 12 December 2016

Published: 23 December 2016

Citation:

Berthier ML and Moreno-Torres I
(2016) Commentary: Visual Feedback
of Tongue Movement for Novel
Speech Sound Learning.
Front. Hum. Neurosci. 10:662.
doi: 10.3389/fnhum.2016.00662

K & M's study represents a step forward in the use of multimodal information for second language learning in healthy subjects and for treating abnormal speech in different conditions (stuttering, apraxia of speech, foreign accent syndrome—FAS-).

K & M also outlined the pathways providing external visual feedback and internal feedback using the neurocomputational ACT model of speech production and perception (Kröger and Kannampuzha, 2008). The key role of cortical areas mediating body awareness (insula) and reward during behavioral training (lateral premotor cortex) has also been discussed. In a recent study using real-time functional magnetic resonance imaging (rtfMRI) during visual neurofeedback, Ninaus et al. (2013) found activation of bilateral anterior insular cortex (AIC), anterior cingulate cortex (ACC), and supplementary motor and dorsomedial and lateral prefrontal areas. Gaining further knowledge on the multiple components of these networks is particularly telling for intensive learning training and for therapeutic purposes because using rtfMRI during neurofeedback humans can learn to voluntarily self-regulate brain activity in circumscribed cortical areas (Caria et al., 2007; Rota et al., 2009). The AIC is one such key circumscribed area susceptible to be voluntarily self-controlled (Ninaus et al., 2013). By acting as a multimodal integration hub, the AIC is in an ideal position for coordinating the activity of several networks devoted to modulate multisensory information (visual, auditory-motor, tactile, somatosensory, Ackermann and Riecker, 2010; Moreno-Torres et al., 2013) involved in feedback of tongue movements during the speech production. In addition, note that healthy subjects activate both AIC during overt speech (Ackermann and Riecker, 2010) and that the AIC together with the left frontal operculum are co-activated during learning

foreign sounds (Ventura-Campos et al., 2013). The AIC also contains the sensory-motor maps that code the subjective feeling of our own movements (body awareness; Critchley et al., 2004; Craig, 2009). Acting conjointly with the ACC, the AIC is engaged in speech initiation (Goldberg, 1985), cognitive control, goal-directed attention, and error detection (Nelson et al., 2010) during the execution of effortful tasks (Engström et al., 2015).

Cue detection and attentional control involved in visual feedback probably depend upon the activation of nicotinic acetylcholine receptors (Demeter and Sarter, 2013) in the AIC and ACC (Picard et al., 2013). In support of this view, learning deficits induced by experimentally-induced insular lesions are reverted by stimulation of cholinergic neurotransmission (Russell et al., 1994). Future studies will examine the modulation of neurochemistry with non-invasive brain stimulation (Fiori et al., 2014), drugs (Berthier and Pulvermüller, 2011), or both to ascertain whether the results of novel speech sound learning reported by K & M using visual feedback of tongue movements can be accelerated and maintained in the long-term. Preliminary evidence of the beneficial effect of combining audio-visual feedback training and cholinergic stimulation on reverting FAS has been demonstrated in a single patient (Berthier et al., 2009; Moreno-Torres et al., 2013). Importantly, the almost complete return to native accent in this case of FAS was associated with partial or total normalization of metabolic activity in several regions of the speech production network, especially in the left AIC and ACC (Figure 1).

In summary, the results of K & M open new important avenues for future research on the role of visual feedback in learning new phonemes. Further refinement of these methodologies coupled with a better understanding of the

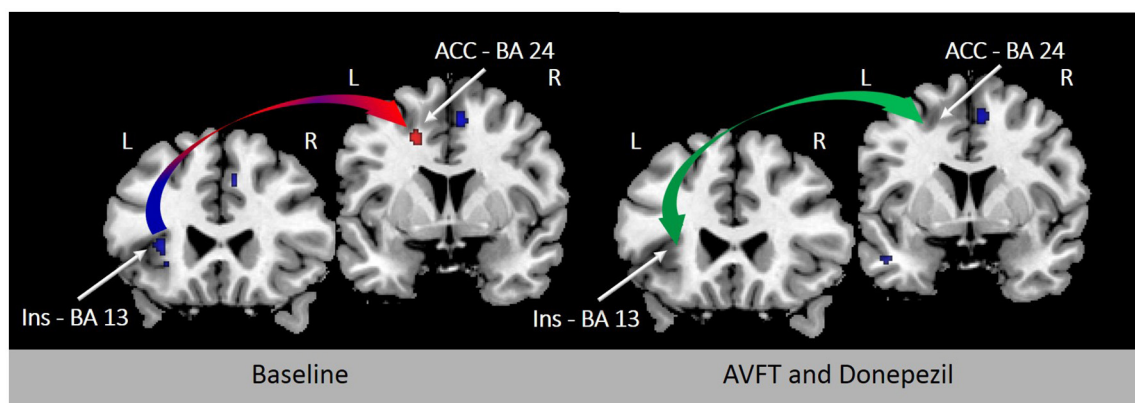


FIGURE 1 | Resting [^{18}F]-fluorodeoxyglucose positron emission tomography (^{18}FDG -PET) before and after treatment in a 44-year-old woman with chronic post-stroke foreign accent syndrome (FAS). Rates of metabolic activity in this single patient were compared with 18 healthy control subjects (female/male: 11/7; mean age \pm SD: 56.6 ± 5.7 years; age range: 47–60 years). The left panel shows T1-weighted coronal sections depicting significant hypometabolism in the left deep frontal operculum and dorsal anterior insula (Talairach and Tournoux peak coordinates: $x = -36$ $y = 14$ $z = 16$) (blue arrow) with significantly increased compensatory metabolic activity in the left ACC (Talairach and Tournoux peak coordinates: $x = -18$ $y = 8$ $z = 40$) (red arrow) before treatment (Moreno-Torres et al., 2013). Clusters in baseline ^{18}FDG -PET were significantly larger with family-wise error correction ($p < 0.05$) in comparison to healthy controls. Deficient phonetic error awareness and monitoring responsible from FAS in this patient were treated with audiovisual feedback training (visually-guided using Praat—(Boersma and Weenink, 2010)—and adjusting patient's emissions to F0 contours of sentence-models) alone and in combination with a cholinergic drug (donepezil, 5 mg/day). An almost complete return of accent to its premorbid characteristics was observed after combined treatment and these changes correlated with normalization of metabolic activity in the left AIC and ACC (green arrow; right panel; Berthier et al., 2009). BA indicates Brodmann area.

neural and chemical mechanisms implicated in tuning tongue movements with audio-visual feedback to learn new phonemes represent a key area of enquiry in the neuroscience of speech.

AUTHOR CONTRIBUTIONS

The two authors (MB and IM) have made substantial, direct and intellectual contribution to the work, and approved it for

publication. MB and IM drafted the article and revised it critically for important intellectual content.

FUNDING

This research was partly supported by a grant from the Spanish Ministerio de Economía e Innovación to the second author (FFI2015-68498P).

REFERENCES

- Ackermann, H., and Riecker, A. (2010). The contribution(s) of the insula to speech production: a review of the clinical and functional imaging literature. *Brain Struct. Funct.* 214, 419–433. doi: 10.1007/s00429-010-0257-x
- Berthier, M. L., Moreno-Torres, I., Green, C., Cid, M. M., García-Casares, N., Juárez, R., et al. (2009). Cholinergic modulation of fronto-insular networks with donepezil and linguistic-emotional imitation therapy in foreign accent syndrome (abstract in Spanish). *Neurología* 24, 623.
- Berthier, M. L., and Pulvermüller, F. (2011). Neuroscience insights improve neurorehabilitation of poststroke aphasia. *Nat. Rev. Neurol.* 7, 86–97. doi: 10.1038/nrneurol.2010.201
- Boersma, P., and Weenink, D. (2010). *Praat: Doing Phonetics by Computer (Version 5.1.25) [Computer Program]*. Available online at: <http://www.praat.org/>
- Caria, A., Veit, R., Sitaram, R., Lotze, M., Weiskopf, N., Grodd, W., et al. (2007). Regulation of anterior insular cortex activity using real-time fMRI. *Neuroimage* 35, 1238–1246. doi: 10.1016/j.neuroimage.2007.01.018
- Craig, A. D. (2009). How do you feel—now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* 10, 59–70. doi: 10.1038/nrn2555
- Critchley, H. D., Wiens, S., Rotshtein, P., Ohman, A., and Dolan, R. J. (2004). Neural systems supporting interoceptive awareness. *Nat. Neurosci.* 7, 189–195. doi: 10.1038/nn1176
- Demeter, E., and Sarter, M. (2013). Leveraging the cortical cholinergic system to enhance attention. *Neuropharmacology* 64, 294–304. doi: 10.1016/j.neuropharm.2012.06.060
- Engström, M., Karlsson, T., Landtblom, A. M., and Craig, A. D. (2015). Evidence of conjoint activation of the anterior insular and cingulate cortices during effortful tasks. *Front. Hum. Neurosci.* 8:1071. doi: 10.3389/fnhum.2014.01071
- Fiori, V., Cipollari, S., Caltagirone, C., and Marangolo, P. (2014). If two witches would watch two watches, which witch would watch which watch? tDCS over the left frontal region modulates tongue twister repetition in healthy subjects. *Neuroscience* 256, 195–200. doi: 10.1016/j.neuroscience.2013.10.048
- Goldberg, G. (1985). Supplementary motor area structure and function: review and hypothesis. *Behav. Brain Sci.* 8, 567–616.
- Katz, W. F., and Mehta, S. (2015). Visual feedback of tongue movement for novel speech sound learning. *Front. Hum. Neurosci.* 9:612. doi: 10.3389/fnhum.2015.00612
- Katz, W., Campbell, T., Wang, J., Farrar, E., Eubanks, J. C., Balasubramanian, A., et al. (2014). “Opti-Speech: a real-time, 3D visual feedback system for speech training,” in *INTERSPEECH* (Singapore), 1174–1178.
- Kröger, B. J., and Kannampuzha, J. (2008). “A neurofunctional model of speech production including aspects of auditory and audio-visual speech perception,” in *International Conference of Auditory-Visual Speech Processing* (Queensland), 83–88.
- Moreno-Torres, I., Berthier, M. L., Del Mar Cid, M., Green, C., Gutiérrez, A., García-Casares, N., et al. (2013). Foreign accent syndrome: a multimodal evaluation in the search of neuroscience-driven treatments. *Neuropsychologia* 51, 520–537. doi: 10.1016/j.neuropsychologia.2012.11.010
- Nelson, S. M., Dosenbach, N. U., Cohen, A. L., Wheeler, M. E., Schlaggar, B. L., and Petersen, S. E. (2010). Role of the anterior insula in task-level control and focal attention. *Brain Struct. Funct.* 214, 669–680. doi: 10.1007/s00429-010-0260-2
- Ninaus, M., Kober, S. E., Witte, M., Koschutnig, K., Stangl, M., Neuper, C., et al. (2013). Neural substrates of cognitive control under the belief of getting neurofeedback training. *Front. Hum. Neurosci.* 7:914. doi: 10.3389/fnhum.2013.00914
- Picard, F., Sadaghiani, S., Leroy, C., Courvoisier, D. S., Maroy, R., and Botlaender, M. (2013). High density of nicotinic receptors in the cingulo-insular network. *Neuroimage* 79, 42–51. doi: 10.1016/j.neuroimage.2013.04.074
- Rota, G., Sitaram, R., Veit, R., Erb, M., Weiskopf, N., Dogil, G., et al. (2009). Self-regulation of regional cortical activity using real-time fMRI: the right inferior frontal gyrus and linguistic processing. *Hum. Brain Mapp.* 30, 1605–1614. doi: 10.1002/hbm.20621
- Russell, R. W., Escobar, M. L., Booth, R. A., and Bermúdez-Rattoni, F. (1994). Accelerating behavioral recovery after cortical lesions. II. *In vivo* evidence for cholinergic involvement. *Behav. Neural Biol.* 61, 81–92.
- Sumby, W. H., and Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *J. Acoust. Soc. Am.* 26, 212–215.
- Ventura-Campos, N., Sanjuán, A., González, J., Palomar-García, A. M., Rodríguez-Pujadas, A., Sebastián-Gallés, N., et al. (2013). Spontaneous brain activity predicts learning ability of foreign sounds. *J. Neurosci.* 33, 9295–9305. doi: 10.1523/JNEUROSCI.4655-12.2013

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Berthier and Moreno-Torres. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.