

**Authors:**

Uri Hasson, Daniel A. Llano, & Gabriele Miceli, & Anthony S. Dick.

**Words Counts:**

ABSTRACT, 56

MAIN TEXT, 996

REFERENCES, 635

ENTIRE TEXT (TOTAL (main text, abstract, references) + ADDRESSES etc.) 1847

**Title:**

Does it talk the talk? On the role of basal ganglia in emotive speech processing

**Author details:**

1. Uri Hasson, Ph.D (Corresponding Author)  
Center for Mind/Brain Sciences (CIMEC) and  
Department of Psychology and Cognitive Science.  
University of Trento.  
Via Delle Regole 101, Mattarello (TN). Italy.  
Phone: +39-0461-882777  
Email: [uri.hasson@unitn.it](mailto:uri.hasson@unitn.it)  
WWW: <http://www.hasson.org>
2. Daniel A. Llano, M.D, Ph.D (co-author)  
The School of Molecular and Cellular Biology.  
University of Illinois at Urbana-Champaign.  
405 North Mathews Avenue. Urbana, IL 61801. USA  
Email: [d-llano@illinois.edu](mailto:d-llano@illinois.edu)  
WWW: <http://mcb.illinois.edu/faculty/profile/d-llano/>
3. Gabriele Miceli, M.D, Ph.D. (co-author)  
Center for Mind/Brain Sciences (CIMEC) and  
Department of Psychology and Cognitive Science.  
University of Trento.  
Via Matteo del Ben, 5/b 38068 Rovereto (TN). Italy  
Email: [gabriele.miceli@unitn.it](mailto:gabriele.miceli@unitn.it)  
WWW: <http://www.unitn.it/en/cimec/11706/gabriele-miceli>
4. Anthony Steven Dick, Ph.D (co-author)  
Department of Psychology.  
Florida International University.  
11200 S.W. 8th Street. Miami, FL 33199. USA  
Email: [adick@fiu.edu](mailto:adick@fiu.edu)  
WWW: <http://faculty.fiu.edu/~adick>

**Abstract:** Ackermann et al.'s phylogenetic account of speech argues that the basal ganglia imbue speech with emotive content. However, a body of work on auditory/emotive processing is inconsistent with attributing this function exclusively to these structures. The account further overlooks the possibility that the emotion-

integration function may be at least in part mediated by the cortico-ponto-cerebellar system.

**Main text:** Ackermann et al.'s phylogenetic account of speech development hinges, in part, on premises related to the role of basal ganglia (BG) in adult human speech production. It argues that in adults, BG imbues speech with emotive content. While the model targets an important and neglected issue, we argue it suffers from two structural weaknesses: On the one hand, it does not sufficiently consider studies of the role of BG in auditory and emotive processing such as those showing that BG damage does not disrupt emotive processing in speech. The argument also overlooks the possibility that the role attributed to the BG may be at least in part mediated by a different system – the cortico-ponto-cerebellar system. We believe the authors' account would be much strengthened if they address these points, which we detail in turn.

*Viability of BG as a speech/emotion synthesizer.* A principle incorporated in contemporary models of speech production is that production occurs under one or more levels of feedback, where potential production errors are monitored either after utterance production (sensory feedback) or prior to it (via internal models; e.g., Hickok, 2012). The authors do not couch their account in an existing speech-production model and leave the issue of feedback underspecified. Nonetheless, if the BG were responsible for imbuing speech with emotive content, it would be expected to have the capacity to monitor and correct for related errors; i.e., evaluate that the intended emotive tone/prosody was instantiated. However, the BG is a weak candidate for such a function. The authors ignore studies indicating (i) that the auditory response in BG is temporally insufficient to provide feedback (Langers & Melcher, 2011) and that it has limited functional connectivity with areas of the temporal cortex mediating language processing (Choi, Yeo, & Buckner, 2012), (ii) that emotive speech processing is mediated mainly by lateral temporal systems while excluding the BG (Kotz, Kalberlah, Bahlmann, Friederici, & Haynes, 2013; Wildgruber, Ackermann, Kreifelts, & Ethofer, 2006), and most importantly (iii) that individuals with BG infarcts are equally sensitive to emotional speech variations as control populations (Paulmann, Ott, & Kotz, 2011; Paulmann, Pell, & Kotz, 2008). These three points argue against the authors' claim that adding prosody to speech depends on integrity of striatum.

The suggested account relies on two additional premises that are not strongly supported by the literature. The first: in adults, the BG can afford coding for emotion since adult perisylvian regions code for syllable motor programs, independently of the BG. Empirical support for this point is tenuous at best: studies using manipulations of syllable frequency have either reported null results (Brendel et al., 2011; Riecker, Brendel, Ziegler, Erb, & Ackermann, 2008) or documented effects in the anterior insula (Carreiras, Mechelli, & Price, 2006). The second: that the BG can merge emotional content due to cross talk between cortico-striatal-thalamic circuits. Although there is anatomical evidence for cross-talk across BG circuits in animal models (Haber, 2003), the functional significance of these needs to be fleshed out.

*On the consideration of alternatives.* A BG-oriented account should address questions such as those above, and equally importantly argue why the BG is the strongest neurobiological candidate for mediating the function in question. The authors do not make such an argument, which is unfortunate since much of the neurobiological argument made here for BG could be made effectively for other structures, such as the cerebellum.

The involvement of the cerebellum in emotional processing is well established. It is implicated in self-generation of various emotional states (Damasio et al., 2000), with different emotions evoking distinct activity patterns in the structure (Baumann & Mattingley, 2012). Damage to the cerebellum affects emotional processing. In animal models, early cerebellar lesions can lead to disrupted emotional processing (Bobee, Mariette, Tremblay-Leveau, & Caston, 2000), and in human adults the Cerebellar Cognitive Affective Syndrome (CCAS; Schmahmann & Sherman, 1998) is a recognized clinical entity associated with blunting of affect. This latter syndrome has been attributed to damage to the posterior vermis, which reduces the cerebellar contribution to perisylvian cortical areas via its outflow to the ventral tier thalamic nuclei (Stoodley & Schmahmann, 2010).

Arguments used by Ackermann et al. in support of their BG hypothesis could also be applied to the cerebellum. For example, FOXP2 expression is found in the cerebellum as well as the caudate (Lai, Gerrelli, Monaco, Fisher, & Copp, 2003; Watkins et al., 2002), and as shown by Ackermann et al. (1992), cerebellar lesions are associated with dysarthria. In addition, activity in the cerebellum, but not BG, discriminates emotive aspects of speech (Kotz et al., 2013). Furthermore, the cerebellum has the capacity for generating an internal forward model of motor-to-auditory predictions of the sort needed to evaluate whether the intended emotive aspect has been communicated (Knolle, Schroger, & Kotz, 2013). While there is no direct examination of this issue for BG, work on motor control suggests that functionally, BG may implement open- rather than closed-loop control of motor actions (Gabrieli, Stebbins, Singh, Willingham, & Goetz, 1997).

It is important to point out that these explanations are not mutually exclusive. Cerebellar and BG circuits involved with language converge at the ventral anterior nucleus of the thalamus, which has also been implicated in language, and can serve as a nidus for cortical feedback via cortico-thalamic projections (Crosson, 2013). Further, cerebellar outflow can directly influence the BG and vice versa, (Bostan, Dum, & Strick, 2013), suggesting that attributing the emotional content of speech to either of these two systems in isolation may not be possible. Given this connectivity, it may be that the cerebellum drives emotion-carrying vocalizations by involving BG, or that the BG trigger emotional behavior that is ultimately modulated by the cerebellum, as would be consistent with a CCAS syndrome. However, data on this issue are lacking.

Summary: Arguing that the BG can imbue speech with emotional content is a significant claim and as such requires additional evidence, accompanied by careful consideration of alternative accounts. We hope this commentary will result in more detailed examination of the aforementioned issues.

## References

- Ackermann, H., Vogel, M., Petersen, D., & Poremba, M. (1992). Speech deficits in ischaemic cerebellar lesions. *J Neurol*, *239*(4), 223-227.
- Baumann, O., & Mattingley, J. B. (2012). Functional topography of primary emotion processing in the human cerebellum. *Neuroimage*, *61*(4), 805-811. doi: 10.1016/j.neuroimage.2012.03.044
- Bobee, S., Mariette, E., Tremblay-Leveau, H., & Caston, J. (2000). Effects of early midline cerebellar lesion on cognitive and emotional functions in the rat. *Behav Brain Res*, *112*(1-2), 107-117.
- Bostan, A. C., Dum, R. P., & Strick, P. L. (2013). Cerebellar networks with the cerebral cortex and basal ganglia. *Trends Cogn Sci*, *17*(5), 241-254. doi: 10.1016/j.tics.2013.03.003
- Brendel, B., Erb, M., Riecker, A., Grodd, W., Ackermann, H., & Ziegler, W. (2011). Do we have a "mental syllabary" in the brain? An fMRI study. *Motor Control*, *15*(1), 34-51.
- Carreiras, M., Mechelli, A., & Price, C. J. (2006). Effect of word and syllable frequency on activation during lexical decision and reading aloud. *Hum Brain Mapp*, *27*(12), 963-972. doi: 10.1002/hbm.20236
- Choi, E. Y., Yeo, B. T., & Buckner, R. L. (2012). The organization of the human striatum estimated by intrinsic functional connectivity. *J Neurophysiol*, *108*(8), 2242-2263. doi: 10.1152/jn.00270.2012
- Crosson, B. (2013). Thalamic mechanisms in language: a reconsideration based on recent findings and concepts. *Brain Lang*, *126*(1), 73-88. doi: 10.1016/j.bandl.2012.06.011
- Damasio, A. R., Grabowski, T. J., Bechara, A., Damasio, H., Ponto, L. L., Parvizi, J., & Hichwa, R. D. (2000). Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nat Neurosci*, *3*(10), 1049-1056. doi: 10.1038/79871
- Gabrieli, J. D., Stebbins, G. T., Singh, J., Willingham, D. B., & Goetz, C. G. (1997). Intact mirror-tracing and impaired rotary-pursuit skill learning in patients with Huntington's disease: evidence for dissociable memory systems in skill learning. *Neuropsychology*, *11*(2), 272-281.
- Haber, S. N. (2003). The primate basal ganglia: parallel and integrative networks. *J Chem Neuroanat*, *26*(4), 317-330.
- Hickok, G. (2012). Computational neuroanatomy of speech production. *Nat Rev Neurosci*, *13*(2), 135-145. doi: 10.1038/nrn3158
- Knolle, F., Schroger, E., & Kotz, S. A. (2013). Cerebellar contribution to the prediction of self-initiated sounds. *Cortex*, *49*(9), 2449-2461. doi: 10.1016/j.cortex.2012.12.012
- Kotz, S. A., Kalberlah, C., Bahlmann, J., Friederici, A. D., & Haynes, J. D. (2013). Predicting vocal emotion expressions from the human brain. *Hum Brain Mapp*, *34*(8), 1971-1981. doi: 10.1002/hbm.22041
- Lai, C. S., Gerrelli, D., Monaco, A. P., Fisher, S. E., & Copp, A. J. (2003). FOXP2 expression during brain development coincides with adult sites of pathology in a

- severe speech and language disorder. *Brain*, 126(Pt 11), 2455-2462. doi: 10.1093/brain/awg247
- Langers, D. R., & Melcher, J. R. (2011). Hearing without listening: functional connectivity reveals the engagement of multiple nonauditory networks during basic sound processing. *Brain Connect*, 1(3), 233-244. doi: 10.1089/brain.2011.0023
- Paulmann, S., Ott, D. V., & Kotz, S. A. (2011). Emotional speech perception unfolding in time: the role of the basal ganglia. *PLoS One*, 6(3), e17694. doi: 10.1371/journal.pone.0017694
- Paulmann, S., Pell, M. D., & Kotz, S. A. (2008). Functional contributions of the basal ganglia to emotional prosody: evidence from ERPs. *Brain Res*, 1217, 171-178. doi: 10.1016/j.brainres.2008.04.032
- Riecker, A., Brendel, B., Ziegler, W., Erb, M., & Ackermann, H. (2008). The influence of syllable onset complexity and syllable frequency on speech motor control. *Brain Lang*, 107(2), 102-113. doi: 10.1016/j.bandl.2008.01.008
- Schmahmann, J. D., & Sherman, J. C. (1998). The cerebellar cognitive affective syndrome. *Brain*, 121 ( Pt 4), 561-579.
- Stoodley, C. J., & Schmahmann, J. D. (2010). Evidence for topographic organization in the cerebellum of motor control versus cognitive and affective processing. *Cortex*, 46(7), 831-844. doi: 10.1016/j.cortex.2009.11.008
- Watkins, K. E., Vargha-Khadem, F., Ashburner, J., Passingham, R. E., Connelly, A., Friston, K. J., . . . Gadian, D. G. (2002). MRI analysis of an inherited speech and language disorder: structural brain abnormalities. *Brain*, 125(Pt 3), 465-478.
- Wildgruber, D., Ackermann, H., Kreifelts, B., & Ethofer, T. (2006). Cerebral processing of linguistic and emotional prosody: fMRI studies. *Prog Brain Res*, 156, 249-268. doi: 10.1016/S0079-6123(06)56013-3