Brief Communications

Enhanced Procedural Learning of Speech Sound Categories in a Genetic Variant of *FOXP2*

Bharath Chandrasekaran,[1,2,3,4,5] Han-Gyol Yi,[1,2] Nathaniel J. Blanco,[2] John E. McGeary,[7,8,9] and W. Todd Maddox[2,4,5,6]

[1]Department of Communication Sciences and Disorders, [2]Department of Psychology, [3]Department of Linguistics, [4]Institute for Neuroscience, [5]Institute for Mental Health Research, and [6]Center for Perceptual Systems, The University of Texas at Austin, Austin, Texas 78712, [7]Providence Veterans Affairs Medical Center, Providence, Rhode Island 02908, [8]Department of Psychiatry and Human Behavior, Brown University, Providence, Rhode Island 02912, and [9]Division of Behavioral Genetics, Rhode Island Hospital, Providence, Rhode Island 02903

A mutation of the forkhead box protein P2 (*FOXP2*) gene is associated with severe deficits in human speech and language acquisition. In rodents, the humanized form of *FOXP2* promotes faster switching from declarative to procedural learning strategies when the two learning systems compete. Here, we examined a polymorphism of *FOXP2* (rs6980093) in humans (214 adults; 111 females) for associations with non-native speech category learning success. Neurocomputational modeling results showed that individuals with the GG genotype shifted faster to procedural learning strategies, which are optimal for the task. These findings support an adaptive role for the *FOXP2* gene in modulating the function of neural learning systems that have a direct bearing on human speech category learning.

Key words: category learning; corticostriatal systems; *FOXP2*; genetic variation; procedural learning; speech

Introduction

A rare point mutation in the forkhead box protein P2 (*FOXP2*) gene is linked to severe neurodevelopmental deficits in speech, language, and orofacial function (Lai et al., 2001). This discovery fueled the examination of *FOXP2* as a candidate gene associated with human speech and language (Fisher and Scharff, 2009). Although the neurobiological mechanisms linking *FOXP2* to human speech and language are an ongoing topic of research, animal models implicate a role for *FOXP2* in modulating the function of corticostriatal systems broadly involved in sensory–motor learning (Enard, 2011). For example, *FoxP2* mutations impair motor skill and auditory motor learning in mice (Kurt et al., 2012). In mice, a humanized form of *FoxP2* induces substantial changes to the microstructure and plasticity of medium-spiny neurons in the striatum (Enard et al., 2009). Behaviorally, these mice show an advantage in learning stimulus–response associations relative to controls when place-based (comparable to a declarative strategy in humans) and response-based (comparable to

a procedural strategy in humans) strategies compete for control (Schreiweis et al., 2014). Given that human language learning is also hypothesized to be mediated by declarative and procedural systems, Schreiweis et al. (2014) posit that *FOXP2* may have affected language evolution by tuning corticostriatal systems differentially, leading to faster proceduralization of complex learned sequences.

Here, we examine the extent to which *FOXP2* variation is associated with the learning of non-native speech sound categories in adults. Learning new speech categories in adulthood is considered a difficult task, characterized by large interindividual differences in learning success (Golestani et al., 2002; Wang et al., 2003; Myers and Swan, 2012). Speech categories are differentiated by multiple, difficult-to-verbalize dimensions. Recent studies posit that speech categories are learned via multiple, competitive, corticostriatal learning systems, with the balance of control between systems relating to individual differences in successful learning (Lim et al., 2014; Yi et al., 2014). Two competitive corticostriatal learning systems have been identified: a prefrontal cortex-mediated reflective system that uses declarative rules to learn new categories and a striatally mediated, reflexive, procedural-based system wherein categorization is not under conscious control (Ashby and Maddox, 2011). During reflexive learning, medium-spiny neurons in the putamen associate a group of sensory units with a motor response when learning is rewarded (Ashby and Maddox, 2011). These systems compete for control during category learning. Consistent with this framework, behavioral and neuroimaging studies show that speech category learning initially engages the reflective system, but is optimally learned by switching to the reflexive, procedural learning system (Yi et al., 2014).

Given the modulatory role of *foxP2* on corticostriatal circuitry and learning in animal models, we hypothesized that genetic vari-

Table 1. Demographic information and cognitive scores

FOXP2	n	Age (y)	Operational span	Music: proficiency	Music: years of training
AA	101 (59)	25.24 (4.51)	41.38 (15.57)	4.98 (2.34)	8.03 (5.83)
AG	72 (36)	25.17 (4.38)	45.31 (15.54)	5.12 (2.05)	9.24 (5.86)
GG	31 (16)	24.35 (3.89)	43.77 (17.69)	5.36 (2.57)	7.75 (7.31)
$p(H_0)$	0.518	0.604	0.306	0.789	0.518

Demographic and cognitive scores across genotypes. Values in parentheses denote SDs, except for those in *n* column, which denote the number of female participants. *p*-values correspond to the null hypotheses that these values are equivalent across all three groups. For *n*, the *p*-value is derived from a χ^2 test on the number of females and males across genotype groups. For all other variables, *p*-values are derived from three-way ANOVAs.

ation in *FOXP2* would predict individual differences in successful learning of novel speech categories. We identified a *FOXP2* single nucleotide polymorphism (SNP rs6980093) that has been shown previously to selectively modulate PFC activity (AA homozygotes > GG homozygotes) during speech processing (Pinel et al., 2012). We hypothesized that the enhanced PFC activity in AA homozygotes may interfere with the transfer of control from the declarative, reflective system to the procedural-based reflexive learning system, resulting in poorer learning in AA homozygotes relative to GG homozygotes. We used neurocomputational modeling to evaluate the effect of *FOXP2* variation on learning strategies. The modeling analyses allowed the examination of learning strategies, thus providing information beyond simple accuracy measures. Our models included declarative, reflective learning strategies, as well as a reflexive learning strategy based on medium-spiny neuron function within the procedural learning system. Consistent with predictions from animal models, we hypothesized that *FOXP2* variation would influence the use of learning strategies and influence the transfer of control to the procedural-based reflexive learning system. Specifically, we predicted a faster shift to procedural learning computational strategies in GG homozygotes relative to AA homozygotes.

Materials and Methods

Participants. Adult English speakers (*N* = 216; 120 females) from ages 18–35 (mean = 25.02; SD = 4.33) with no prior experience with a tone language were recruited from the Austin, Texas, community. Participants reported no prior history of speech, language, neurological, or hearing disorders. Participants were excluded if they reported a history of Axis I DSM-IV disorders using the Mini International Neuropsychiatric Interview (Sheehan et al., 1998). Participants were also excluded if they had participated in a previous speech category learning study (*n* = 12). All participants were genotyped for a polymorphism in *FOXP2* (SNP rs6980093): AA (*n* = 101; 59 females), AG (*n* = 72; 36 females), and GG (*n* = 31; 16 females). Procedures were approved by the Institutional Review Board at the University of Texas at Austin. Participants underwent the automated operation span (OSPAN) test (Unsworth et al., 2005), which measures working memory and executive attention. In this task, participants recalled letters with simple mathematical problems presented as a distractor. Across genotypes, participants did not significantly differ in age, sex, number of years of musical training, self-rating of musicianship on a 1 (low) to 10 (high musical ability), or OSPAN scores (Table 1). In addition, these factors and genotypic variations were found not to interact in affecting speech category learning performance at a $\alpha <$ 0.05 criterion in mixed-effects modeling analyses.

Speech learning task stimuli. Two native Mandarin Chinese speakers (1 female) produced four Mandarin Chinese lexical tones in citation form. These tones are characterized by fundamental frequency (F0) height and slope variations such as: high-level, low-rising, low-dipping, and high-falling (Fig. 1*a*). The lexical tones were produced in the context of five monosyllabic words: /bu/, /di/, /lu/, /ma/, and /mi/. Stimuli were normalized for RMS amplitude at 70 dB and duration of 0.44 s.
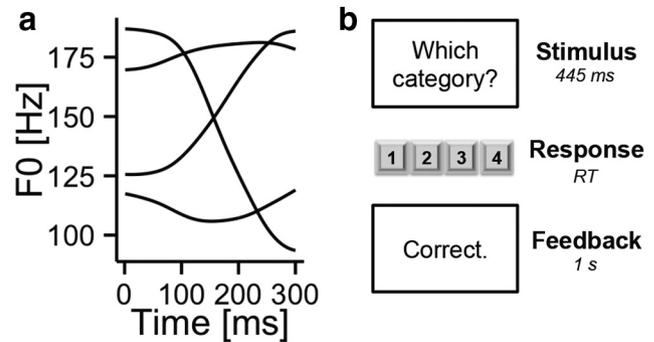

Figure 1. *a*, Example fundamental frequency (F0, a correlate of pitch) patterns for the four Mandarin lexical tone categories produced by a speaker in the context of syllable /di/: high-level, low-rising, low-dipping, and high-falling. *b*, Experimental procedure for the speech category learning task.

Speech learning task procedures. Participants performed a speech category learning task in a sound-attenuated booth. On each trial, a speech sound was presented via headphones. Participants were instructed to categorize each sound into 1 of 4 categories using number keys (1, 2, 3, or 4) on a keyboard. No further instructions were given. Feedback was immediately presented after the response according to whether a correct categorization was made (Fig. 1*b*). Therefore, participants learned to map sounds to specific categories via a motor response, using feedback to monitor errors. Each of the speech sounds (*n* = 40) was presented once within each of the five learning blocks using a randomized sequence.

Genotyping procedures. The SNP in the *FOXP2* was assayed using the Taqman genotyping assay C__11437347_10 (Applied Biosystems) using an ABI 7900HT Real-Time PCR System. The frequency of the *FOXP2* (SNP rs6980093) genotypes (*n* = 216; AA: *n* = 103; AG: *n* = 78; GG: *n* = 35) differed from the Hardy–Weinberg equilibrium (χ^2 = 8.496, *p* = 0.004). The deviation could be due to nonrandom mating, natural selection, genetic drift, genotyping error, or sample selection based upon a phenotype that is associated with the tested genotype (Sham, 1998; Hosking et al., 2004).

Accuracy analysis. Participants' response on each trial was coded as "correct" or "incorrect." A mixed-logit analysis was conducted to estimate the log odds of a correct response (Bates D, Maechler M, Bolker B, lme4: Linear mixed-effects models using S4 classes software). The model included a fixed effect of genotype (AA, AG, and GG) corrected for random intercepts for participant and trial number.

Theoretical and neurocomputational modeling approach. The neurocomputational models examined here reflect the neurobiological properties of the reflective and reflexive corticostriatal learning systems. (Nomura et al., 2007; Seger and Miller, 2010). The reflective learning system involves the executive corticostriatal loop mediated by the dorsolateral prefrontal cortex, head of the caudate nucleus, the anterior cingulate cortex, and the medial temporal lobe. These brain structures are involved in the generation, selection, and maintenance of declarative rules. In contrast, the reflexive learning system involves procedural mapping of an abstract motor response with sensory cells. This sensory–motor mapping requires a reinforcement signal from the ventral striatum (Ashby and Ennis, 2006; Seger, 2008). Modeling was conducted on the response pattern data to evaluate individual participants' response strategies in 40-trial blocks. Each stimulus was reduced to a point in the 2D space defined by F0 height and slope (Fig. 2*a*), corresponding to the critical dimensions underlying tone perception across languages (Chandrasekaran et al., 2007). A previous study showed that native English speakers show a bias toward F0 height and initially separate perceptual spaces based on the sex of the talker. This strategy, when modeled, accounted for more variance than a model that assumes a single perceptual space across talkers (Maddox and Chandrasekaran, 2014). Therefore, stimulus spaces were constructed separately for male and female speakers for modeling purposes. We explored three classes of reflective models (unidimensional F0 height, unidimensional F0 slope, and a multidimensional conjunctive model), a single reflexive model (striatal pattern clas-
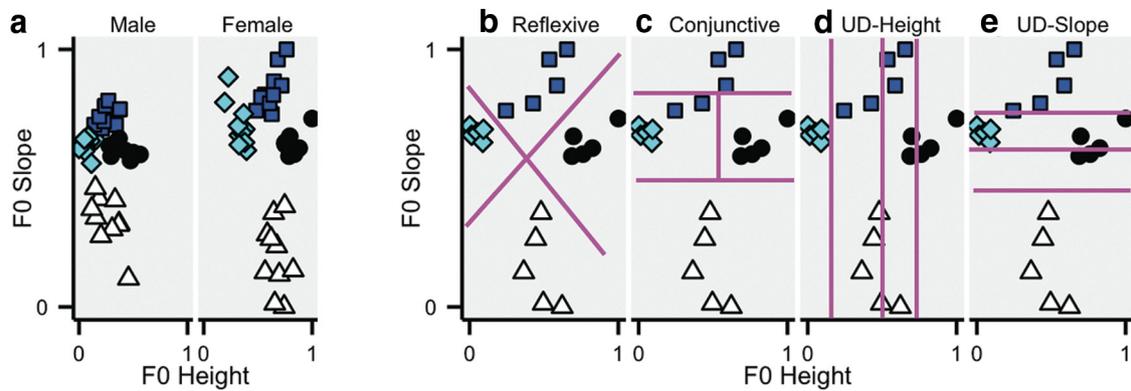
Figure 2. ***a***, Scatterplot of all stimuli from the speech category learning experiment. Two dimensions (F0 height, F0 slope) have been shown to underlie lexical tone perception in previous studies. Scatterplots of the responses, along with the decision boundaries that separate response regions from hypothetical participant, are shown for the striatal pattern classifier (SPC), a reflexive strategy (***b***); conjunctive (***c***), unidimensional height (***d***), and unidimensional slope (***e***) models (***c***–***e*** are considered reflective strategies). Black circles denote the high-level tone, blue squares denote the low-rising tone, the cyan lozenges denote the low-dipping tone, and the white triangles denote the high-falling tone.
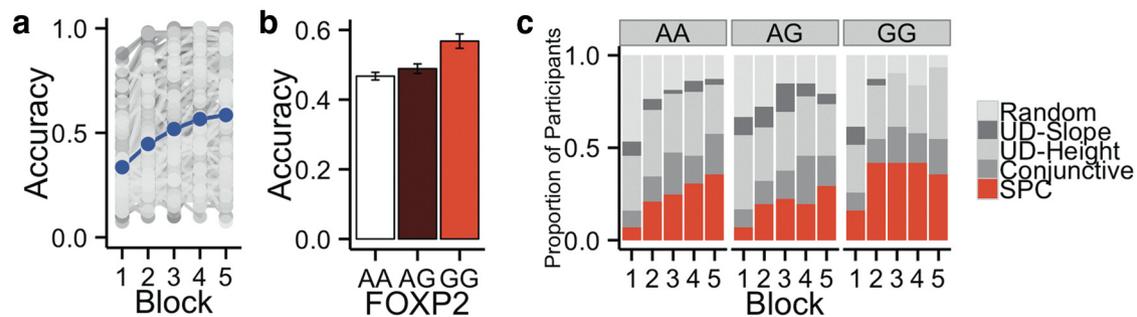

Figure 3. ***a***, Individual performance learning curves. The *x*-axis denotes a block of 40 learning trials. The *y*-axis denotes average accuracy within each block per participant. The blue line denotes the average performance across all participants. Different shades of gray colors denote individual differences in performance. ***b***, Average performance across genotype groups: AA (white), AG (brown), and GG (red). ***c***, Strategy use per block as revealed by neurocomputational modeling separately for carriers of each genotype (AA, AG, GG). Colors denote response strategies (SPC: reflexive; CJ: conjunctive; UD-Height: unidimensional height; UD-Slope: unidimensional slope; Random: random responder). Across all genotypes, reflexive strategy use increases with progression of learning blocks. The GG genotype uses the reflexive strategy earlier relative to AA genotype.

sifier; SPC), and a random responder model. The model parameters were estimated using maximum likelihood procedures (Wickens, 1982) and model fits were compared using Akaike's Information Criterion (Akaike, 1974).

The reflexive learning system was modeled using the SPC computational model (Ashby and Ennis, 2006). The application of this model in the auditory domain is motivated by the many-to-one convergence from the auditory cortex to the striatum (Yeterian and Pandya, 1998). In the SPC model, category learning is instantiated via association of each category label with a cluster of striatal medium-spiny neurons ("striatal units"; Ashby and Ennis, 2006). The SPC assumed a set of four striatal units (one per category) in each of the 2D spaces (see Fig. 2*b*). The SPC can be viewed as a prototype model in which each category is represented by a single striatal unit and stimuli are assigned to categories using a minimum distance rule. Category responses are based on the distance to each striatal unit, with closer categories being more likely. The location of each striatal unit is estimated from a fit to the data. In contrast, the conjunctive and unidimensional models assume that the decision bounds are parallel to the coordinate axes. For example, the conjunctive model outlined in Figure 2*c* assumes that the two verbalizable criteria along the F0 slope dimension are used to separate the stimuli into falling (category 4), rising (category 2), or level (category 1 or 3) slope. Then, a single criterion along the F0 height dimension is used to classify categories 1 (high) and 3 (low). The unidimensional reflective models assume that the participant sets three criteria only along the F0 height (Fig. 2*d*) or F0-slope (Fig. 2*e*) dimension, ignoring the other dimension. The random responder model assumed a fixed probability of responding category 1, 2, 3, or 4 regardless of stimulus properties, allowing for response biases.

First, to test whether the optimal reflexive strategy was used earlier relative to other strategies, a mixed-effects modeling analysis was run with the mean-centered block number as the dependent variable and the strategy as the fixed effect (SPC as the reference), corrected for by-participant random intercepts. Positive estimates would suggest association of each strategy with blocks of the latter half, whereas negative estimates would suggest an association of each strategy with blocks of the earlier half. Second, to assess whether a particular genotype was associated with earlier use of SPC strategies, we ran a mixed-effects modeling analysis with binomial logit link in which the dependent variable was set as per-block SPC use (TRUE vs FALSE) for each participant. The fixed effects included genotype, block number (1–5), and their interaction terms, corrected for by-participant random intercepts.

Results

Accuracy

Participants improved on their ability to learn the tonal categories, from 33% accuracy (SD = 16%) in the first training block to 59% accuracy (SD = 27%) in the final training block (Fig. 3*a*). Consistent with the hypothesis that the GG genotype for *FOXP2* (SNP rs6980093) is associated with enhanced speech category learning, the GG homozygotes were more likely to correctly identify each speech category relative to the AA homozygotes, who served as the reference level, $b = 0.48$, SE $= 0.22$, $z = 2.22$, $p = 0.026$. There was no evidence that the AG heterozygotes learned the speech categories better than the AA homozygotes, $b = 0.12$, SE $= 0.16$, $z = 0.75$, $p = 0.45$. The intercept was not significant,

$b = -0.13$, SE $= 0.11$, $z = -1.22$, $p = 0.22$, suggesting that the accuracy did not deviate from 50% probability among the AA homozygotes. When the same analysis was run with the GG genotype as the reference level, the AG heterozygotes showed a trend to be less likely to correctly identify each speech category relative to the GG homozygotes, $b = -0.36$, SE $= 0.23$, $z = -1.58$, $p = 0.11$, but the difference did not meet the $\alpha < 0.05$ criterion. In addition, we identified non-learners as participants who failed to perform higher than the chance level (25%) at the final block of training. Fifteen of 101 AA homozygotes, 12 of 72 AG heterozygotes, and 3 of 31 GG homozygotes were identified as non-learners. A χ^2 test on these proportions failed to show a statistically significant difference across groups, $\chi^2(2, n = 204) = 0.85$, $p = 0.65$.

Neurocomputational modeling
A mixed-effects analysis was conducted to test whether the reflexive strategy was used later than were the other strategies. The intercept, which modeled the average mean-centered block number associated with occurrence of the reflexive strategy use, was significant, $b = 0.42$, SE $= 0.088$, $t = 4.83$, $p = 1.62 \times 10^{-6}$, indicating that reflexive strategies were more likely to be used during the latter half of learning. Random, unidimensional height, and unidimensional slope strategies were associated with earlier blocks relative to reflexive strategy, $b = -0.97$, SE $= 0.13$, $t = -7.56$, $p = 9.41 \times 10^{-14}$; $b = -0.48$, SE $= 0.12$, $t = -4.09$, $p = 4.65 \times 10^{-5}$; $b = -0.82$, SE $= 0.19$, $t = -4.24$, $p = 2.46 \times 10^{-5}$, respectively. No evidence was found that conjunctive strategy was used earlier or later than was the reflexive strategy, $b = -0.085$, SE $= 0.14$, $t = -0.62$, $p = 0.54$.

Reflexive strategy use across genotypes (using AA as the reference) was assessed with a mixed-effects modeling analysis where the dependent variable noted whether SPC was the best fitting model in each block per participant. The block effect was significant, $b = 0.52$, SE $= 0.095$, $z = 5.49$, $p = 4.04 \times 10^{-8}$, suggesting that the AA homozygotes used the reflexive strategy more toward the later blocks. Relative to the AA homozygotes, the GG homozygotes used the reflexive strategy earlier in learning, $b = 1.66$, SE $= 0.68$, $z = 2.44$, $p = 0.015$. However, they did not differ from the AA homozygotes in how the reflexive strategy use increased later in learning, $b = -0.27$, SE $= 0.17$, $z = -1.59$, $p = 0.11$. The AG heterozygotes did not differ from the AA homozygotes early in learning, $b = 0.075$, SE $= 0.59$, $z = 0.13$, $p = 0.90$, nor in terms of increased reflexive strategy use across the blocks, $b = -0.14$, SE $= 0.15$, $z = -0.93$, $p = 0.36$. The intercept was significant, $b = -3.29$, SE $= 0.40$, $z = -8.28$, $p < 2 \times 10^{-16}$, suggesting that less than 50% of the AA homozygotes initiated learning with a reflexive strategy. Together, these results suggested that the GG genotype was associated with increased reflexive strategy use in the earlier stage of learning relative to the AA genotype (Fig. 3c). Because the conjunctive strategy, which is a reflective strategy, was used at a similar learning stage as the reflexive strategy, an additional statistical analysis (AA as the reference) was conducted to evaluate conjunctive strategy use across genotypes. AA homozygotes were more likely to use conjunctive strategy toward the later learning blocks relative to early learning blocks, $b = 0.24$, SE $= 0.096$, $z = 2.50$, $p = 0.012$. This pattern did not differentiate with GG homozygotes, $b = -0.045$, SE $= 0.77$ $z = -0.058$, $p = 0.95$, or with AG heterozygotes, $b = -0.14$, SE $= 0.58$, $z = -0.25$, $p = 0.81$. Notably, no block-by-genotype interactions were found to be significant for GG homozygotes, $b = -0.026$, SE $= 0.20$, $z = -0.13$, $p = 0.90$, or for AG heterozygotes, $b = 0.021$, SE $=$

0.15, $z = 0.14$, $p = 0.89$. To summarize, all three groups were more likely to use conjunctive strategy in later blocks. Therefore, conjunctive strategy use was not modulated by genotype (Fig. 3c).

Discussion
Learning to categorize nonspeech categories is acknowledged to be a challenging task in adulthood. Speech categories are multidimensional and often difficult to learn using declarative rules and thus may require a procedural-based mapping of non-native acoustic patterns to motoric output (Chandrasekaran et al., 2014; Lim et al., 2014; Yi et al., 2014). Here, we show that individual differences in learning to categorize non-native speech sounds successfully are strongly associated with variation in the FOXP2 gene. The variant associated with enhanced speech category learning was also associated with earlier use of procedural-based reflexive strategies.

Our results parallel findings from animal models that demonstrate an impact of FoxP2 on sensory–motor learning. Injection of FoxP2 knockdown virus into Area X, a homolog of the basal ganglia in humans, impairs song learning in zebrafinch "pupils" (Haesler et al., 2007). In the rodent model, FoxP2 mutations impair auditory–motor learning (Kurt et al., 2012). A human-like FoxP2 gene induced in mice models altered plasticity in the medium-spiny neurons within the striatum and induced faster switching from declarative to procedural learning strategies (Enard, 2011; Schreiweis et al., 2014).

Human studies have largely focused on the association between FOXP2 and speech and language processes, a link that came into prominence via a genetic examination of the well studied KE family (Lai et al., 2001). More than half of the members of the KE family showed severe speech and language acquisition deficits. Interestingly, the affected members showed intact pitch perception and production abilities relative to unaffected members (Alcock et al., 2000). In our study, we found that FOXP2 variation in a neurotypical population affects the ability to learn to categorize non-native pitch patterns. We posit that FOXP2 variation does not alter pitch perception per se, but likely modulates the learning processes involved in mapping novel pitch patterns to specific categories. This view is consistent with an emerging view from animal models suggesting a broader role for FOXP2 in mediating the dynamics of learning (Enard, 2011). A faster switch to procedural learning strategies may be useful in automatizing complex motoric behaviors, thereby freeing cortical resources for other cognitive processes (Schreiweis et al., 2014).

We note several limitations of our study. First, we examined a single FOXP2 SNP that had been identified to modulate prefrontal activation during a speech processing task (Pinel et al., 2012). However, this SNP was not found to affect neuroanatomy, so future neuroimaging studies should investigate whether the SNP can affect functional processes without altering brain structure (Hoogman et al., 2014). Second, our study did not examine the impact of FOXP2 on auditory processing ability or speech production. Affected members of the KE family demonstrate significant issues with oral–motor function and speech production, but limb-motor function is unaffected. Because learning to produce novel speech patterns involves a sensory–motor mapping process mediated by the basal ganglia (Zenon and Olivier, 2014), we predict that learning novel speech output may thus also be associated with FOXP2 variation in a neurotypical population. Finally, future studies could use targeted next-generation sequencing to identify all com-

mon and rare variants, conducting cellular assays of *FOXP2* variants to identify potential functional consequences of these SNPs.

In conclusion, we present behavioral and neurocomputational evidence that *FOXP2* genetic variation is associated with non-native speech category learning in adults. The GG genotype showed performance advantages relative to the AA genotype due to a faster switch to a more reflexive, procedural-based learning system. These results offer a starting point in understanding the impact of *FOXP2* in human learning.

References

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Auto Control 19:716–723. CrossRef

Alcock KJ, Passingham RE, Watkins K, Vargha-Khadem F (2000) Pitch and timing abilities in inherited speech and language impairment. Brain Lang 75:34–46. CrossRef Medline

Ashby FG, Ennis JM (2006) The role of the basal ganglia in category learning. In: The psychology of learning and motivation (Ross BH, ed), pp 1–36. New York: Elsevier.

Ashby FG, Maddox WT (2011) Human category learning 2.0. Ann N Y Acad Sci 1224:147–161. CrossRef Medline

Chandrasekaran B, Gandour JT, Krishnan A (2007) Neuroplasticity in the processing of pitch dimensions: A multidimensional scaling analysis of the mismatch negativity. Restor Neurol Neurosci 25:195–210. Medline

Chandrasekaran B, Koslov SR, Maddox WT (2014) Toward a dual-learning systems model of speech category learning. Front Psychol 5:825. CrossRef Medline

Enard W (2011) FOXP2 and the role of cortico-basal ganglia circuits in speech and language evolution. Curr Opin Neurobiol 21:415–424. CrossRef Medline

Enard W, Gehre S, Hammerschmidt K, Hölter SM, Blass T, Somel M, Brückner MK, Schreiweis C, Winter C, Sohr R, Becker L, Wiebe V, Nickel B, Giger T, Müller U, Groszer M, Adler T, Aguilar A, Bolle I, Calzada-Wack J, et al. (2009) A humanized version of *FOXP2* affects cortico-basal ganglia circuits in mice. Cell 137:961–971. CrossRef Medline

Fisher SE, Scharff C (2009) FOXP2 as a molecular window into speech and language. Trends Genet 25:166–177. CrossRef Medline

Golestani N, Paus T, Zatorre RJ (2002) Anatomical correlates of learning novel speech sounds. Neuron 35:997–1010. CrossRef Medline

Haesler S, Rochefort C, Georgi B, Licznerski P, Osten P, Scharff C (2007) Incomplete and inaccurate vocal imitation after knockdown of *FoxP2* in songbird basal ganglia nucleus area X. PLoS Biol 5:e321. CrossRef Medline

Hoogman M, Guadalupe T, Zwiers MP, Klarenbeek P, Francks C, Fisher SE (2014) Assessing the effects of common variation in the *FOXP2* gene on human brain structure. Front Hum Neurosci 8:473. Medline

Hosking L, Lumsden S, Lewis K, Yeo A, McCarthy L, Bansal A, Riley J, Purvis I, Xu CF (2004) Detection of genotyping errors by Hardy–Weinberg equilibrium testing. Eur J Hum Genet 12:395–399. CrossRef Medline

Kurt S, Fisher SE, Ehret G (2012) Foxp2 mutations impair auditory-motor association learning. PLoS One 7:e33130. CrossRef Medline

Lai CS, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP (2001) A forkhead-domain gene is mutated in a severe speech and language disorder. Nature 413:519–523. CrossRef Medline

Lim SJ, Fiez JA, Holt LL (2014) How may the basal ganglia contribute to auditory categorization and speech perception? Front Neurosci 8:230. Medline

Maddox WT, Chandrasekaran B (2014) Tests of a dual-system model of speech category learning. Biling (Camb Engl) 17:709–728. CrossRef Medline

Myers EB, Swan K (2012) Effects of category learning on neural sensitivity to non-native phonetic categories. J Cogn Neurosci 24:1695–1708. CrossRef Medline

Nomura EM, Maddox WT, Filoteo JV, Ing AD, Gitelman DR, Parrish TB, Mesulam MM, Reber PJ (2007) Neural correlates of rule-based and information-integration visual category learning. Cereb Cortex 17:37–43. Medline

Pinel P, Fauchereau F, Moreno A, Barbot A, Lathrop M, Zelenika D, Le Bihan D, Poline JB, Bourgeron T, Dehaene S (2012) Genetic variants of *FOXP2* and KIAA0319/TTRAP/THEM2 locus are associated with altered brain activation in distinct language-related regions. J Neurosci 32:817–825. CrossRef Medline

Schreiweis C, Bornschein U, Burguière E, Kerimoglu C, Schreiter S, Dannemann M, Goyal S, Rea E, French CA, Puliyadi R, Groszer M, Fisher SE, Mundry R, Winter C, Hevers W, Pääbo S, Enard W, Graybiel AM (2014) Humanized *FOXP2* accelerates learning by enhancing transitions from declarative to procedural performance. Proc Natl Acad Sci U S A 111: 14253–14258. CrossRef Medline

Seger CA (2008) How do the basal ganglia contribute to categorization? Their roles in generalization, response selection, and learning via feedback. Neurosci Biobehav Rev 32:265–278. CrossRef Medline

Seger CA, Miller EK (2010) Category learning in the brain. Annu Rev Neurosci 33:203–219. CrossRef Medline

Sham P (1998) Statistics in human genetics. London, UK: Arnold.

Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC (1998) The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 59:22–33; quiz 34–57. Medline

Unsworth N, Heitz RP, Schrock JC, Engle RW (2005) An automated version of the operation span task. Behav Res Methods 37:498–505. CrossRef Medline

Wang Y, Sereno JA, Jongman A, Hirsch J (2003) fMRI evidence for cortical modification during learning of Mandarin lexical tone. J Cogn Neurosci 15:1019–1027. CrossRef Medline

Wickens TD (1982) Models for behavior: stochastic processes in psychology. San Francisco: WH Freeman.

Yeterian EH, Pandya DN (1998) Corticostriatal connections of the superior temporal region in rhesus monkeys. J Comp Neurol 399:384–402. Medline

Yi H, Maddox WT, Mumford JA, Chandrasekaran B (2014) The role of corticostriatal systems in speech category learning. Cereb Cortex. In press.

Zenon A, Olivier E (2014) Contribution of the basal ganglia to spoken language: Is speech production like the other motor skills? Behav Brain Sci 37:576; discussion 577–604. CrossRef Medline