

A hypothesis on improving foreign accents by optimizing variability in vocal learning brain circuits

Anna J. Simmonds^{1*}

¹Division of Brain Sciences, Imperial College London, United Kingdom

Submitted to Journal:
Frontiers in Human Neuroscience

ISSN:
1662-5161

Article type:
Hypothesis & Theory Article

Received on:
18 Jun 2015

Accepted on:
20 Oct 2015

Provisional PDF published on:
20 Oct 2015

Frontiers website link:
www.frontiersin.org

Citation:
Simmonds AJ(2015) A hypothesis on improving foreign accents by optimizing variability in vocal learning brain circuits. *Front. Hum. Neurosci.* 9:606. doi:10.3389/fnhum.2015.00606

Copyright statement:
© 2015 Simmonds. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution and 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.

This Provisional PDF corresponds to the article as it appeared upon acceptance, after peer-review. Fully formatted PDF and full text (HTML) versions will be made available soon.

1 A hypothesis on improving foreign accents by optimizing 2 variability in vocal learning brain circuits

3
4 Anna J Simmonds
5 Division of Brain Sciences, Imperial College London, UK
6

7 **Abstract**

8 Rapid vocal motor learning is observed when acquiring a language in early
9 childhood, or learning to speak another language later in life. Accurate
10 pronunciation is one of the hardest things for late learners to master and they
11 are almost always left with a non-native accent. Here I propose a novel
12 hypothesis that this accent could be improved by optimizing variability in vocal
13 learning brain circuits during learning. Much of the neurobiology of human vocal
14 motor learning has been inferred from studies on songbirds. Jarvis (2004)
15 proposed the hypothesis that as in songbirds there are two pathways in humans:
16 one for learning speech (the striatal vocal learning pathway), and one for
17 production of previously learnt speech (the motor pathway). Learning new
18 motor sequences necessary for accurate non-native pronunciation is challenging
19 and I argue that in late learners of a foreign language the vocal learning pathway
20 becomes inactive prematurely. The motor pathway is engaged once again and
21 learners maintain their original native motor patterns for producing speech,
22 resulting in speaking with a foreign accent. Further, I argue that variability in
23 neural activity within vocal motor circuitry generates vocal variability that
24 supports accurate non-native pronunciation. Recent theoretical and
25 experimental work on motor learning suggests that variability in the motor
26 movement is necessary for the development of expertise. I propose that there is
27 little trial-by-trial variability when using the motor pathway. When using the
28 vocal learning pathway variability gradually increases, reflecting an exploratory
29 phase in which learners try out different ways of pronouncing words, before
30 decreasing and stabilizing once the 'best' performance has been identified. The
31 hypothesis proposed here could be tested using behavioral interventions that
32 optimize variability and engage the vocal learning pathway for longer, with the
33 prediction that this would allow learners to develop new motor patterns that
34 result in more native-like pronunciation.
35

36 **Introduction**

37 *Vocal learning*

38 Vocal learning is the ability to imitate sounds that are heard, as opposed to
39 producing innate vocalizations. Most mammals are not vocal learners and can
40 only produce innate calls that remain unmodified throughout life (Petkov and
41 Jarvis, 2012). Instead they are auditory learners and through experience can
42 readily distinguish environmental sounds, making an appropriate response to
43 what is heard, e.g. a command to 'sit', without the ability to produce it (Jarvis,
44 2004, 2006). In contrast, humans are highly skilled auditory and vocal learners.
45 We are not born with speech and must learn by listening and practicing. Much of

1 the neurobiology of vocal learning has been inferred from studies on songbirds
2 and there are clear anatomical parallels between song learning birds and
3 humans (Figure 1). Humans and songbirds both have a direct projection from
4 motor cortex to motor neurons in the brainstem controlling movements
5 required for vocalizations (larynx in humans and trachea and syrinx in
6 songbirds). This projection is absent in non-learning birds such as chickens, and
7 non-vocal learning primates, such as macaque monkeys (Petkov and Jarvis,
8 2012) (Figure 1). Vocal learning, and motor learning more generally, involves
9 the basal ganglia, which is the focus of the hypothesis presented here. It has been
10 shown that basal ganglia circuitry is involved to a greater extent in motor
11 learning than performance of acquired behaviors (Hikosaka et al., 1999;
12 Hikosaka et al., 2002). There have also been important distinctions made
13 between different regions within the basal ganglia at different stages of motor
14 learning, with the anterior striatum being involved in learning and the posterior
15 striatum in production of overlearned automatic movements (Graybiel, 2008;
16 Jueptner and Weiller, 1998; Miyachi et al., 1997; Yin et al., 2009). The hypothesis
17 presented here focuses on the learning of foreign speech, which requires novel
18 motor movements rather than previously acquired familiar articulatory
19 movements used for native speech.

20

21 *Speech acquisition in infancy*

22 Human infants begin speech acquisition by listening to speech in their
23 environment. They are skilled both in auditory learning, memorizing the
24 communicative sounds of people they interactive with, as well as in vocal
25 learning, from babbling and single word production to articulating well-formed
26 sentences. Stages of speech development start at a universal level and an infant
27 has the ability to learn any language and will start learning the language to which
28 they are exposed. At around 7 months for perception and 10 months for
29 production, speech becomes language-specific. Although infants produce non-
30 speech sounds from birth and vowel-like sounds at around 3 months, canonical
31 babbling does not appear until around 7 months. Language-specific speech
32 production is observed at around 10 months and word production at around a
33 year (Kuhl, 2004; Simmonds et al., 2011b).

34

35 *Speech acquisition later in life*

36 In contrast, when older children and adults begin learning a foreign language,
37 they do not start with a perception phase, a period of listening to language
38 without attempting production of speech sounds. Instead they begin producing
39 speech early on in the learning process, at the same time as undergoing auditory
40 learning. Unlike infants, older learners do not undergo a babbling phase but
41 move straight to word meaning and phrase production, which is influenced by
42 the native language. Using a listening task in bilinguals who learnt a second
43 language after the age of 12, it has been shown that there is a strong tendency to
44 translate a word in a foreign language (L2) into its native (L1) equivalent
45 (Thierry and Wu, 2007; Wu and Thierry, 2010). Similarly, during L2 covert word
46 production, both L2 and L1 phonological representations are retrieved (Wu and
47 Thierry, 2011). Proficient use of vocabulary and grammar are essential skills, but
48 can be learnt instructively, for example from books. However, acquiring a native-
49 like accent requires repeated motor practice, with the accuracy of articulation

1 dependent on repeated attempts to match auditory exemplars of correct
2 pronunciation. Even then, there is considerable inter-individual variability in
3 achieving accurate pronunciation, both in terms of learning strategies and in
4 attainment (Bley-Vroman, 1990) and individual differences in performance have
5 been shown to correlate with structural brain differences (Golestani et al., 2007;
6 Golestani and Pallier, 2006). The challenge of speaking a foreign language is a
7 problem faced by students and teachers of second language education around
8 the world, and pronunciation errors substantially affect communication skills.

9
10 This challenge has effects on both the spoken performance in a foreign language,
11 and the neural systems involved. The 'native-likeness' of an accent, as judged by
12 native speakers, declines over time as the age at which the speaker starts using
13 the foreign language increases. Italian immigrants arriving in the US were
14 deemed to have a native-like accent if they arrived before the age of two,
15 whereas those arriving as teenagers or young adults had accents that clearly
16 marked them as non-native speakers (Flege, 1995). Perhaps one of the most
17 famous examples of a marked foreign accent in a highly proficient user of a
18 foreign language is Józef Teodor Konrad Korzeniowski, better known by his
19 anglicized name, Joseph Conrad. As a late learner of English as a foreign language
20 he mastered the language to such an extent that he was able to produce great
21 works of fiction in English (his *third* language), yet was left with such a thick
22 Polish accent that he was reported to be incomprehensible. Scovel (1988) coined
23 the term the 'Joseph Conrad phenomenon', referring to the mismatch between
24 lexical, morphological and syntactic proficiency, and pronunciation. Even for
25 highly proficient bilinguals, having learnt a language later in life results in
26 differences in activation patterns during speech production. Speaking in a non-
27 native, relative to native, language requires greater engagement of motor-
28 sensory control systems (Simmonds et al., 2011a).

29
30 In addition to age at the time of learning, other factors claimed to affect the
31 degree of foreign accent include gender, amount of time spent in an L2-speaking
32 environment, amount of L1 and L2 use, formal instruction, motivation and
33 language learning aptitude (Piske et al., 2001). Another explanation for the
34 failure to acquire the native accent in a foreign language is that late bilinguals
35 use the same syllable representation for both of their languages, which results in
36 producing non-native L1-like patterns in their L2. In contrast, early bilinguals
37 have separate representations for their two languages, even for syllables that are
38 shared across the languages (Alario et al., 2010). The present paper presents a
39 novel hypothesis on what might explain the persistent accent in late language
40 learners and considers how it could be improved. The hypothesis is informed by
41 findings from vocal learning research in songbirds and motor learning more
42 generally, as well as our previous work particularly focusing on the response of
43 the anterior striatum during adult human vocal learning (Simmonds et al., 2014).
44 Although the anterior striatum was initially active during production of
45 unfamiliar foreign speech, activity in this region rapidly declined. The decline in
46 the striatum happened over the course of the first scanning session, even before
47 formal training. No decline was found for pronunciation of native non-word
48 stimuli, indicating that the reduction was not an effect of novelty. These findings
49 suggest that late language learners do not maintain use of the vocal learning

1 pathway during learning. Although no direct comparison has been made
2 between early and late language learners in terms of activity in the basal ganglia-
3 forebrain-thalamic circuit, a likely finding would be that early learning of a
4 native language would engage this circuit. However, without research on human
5 infants during speech acquisition, this remains speculative.

6 *Parallels between song learning birds and humans for song and speech*

7 As discussed above, humans are highly skilled auditory and vocal learners. Vocal
8 learning also exists in parrots and oscine songbirds (order: Passeriformes)
9 (Mooney, 2009; Petkov and Jarvis, 2012), hummingbirds (Jarvis et al., 2000), and
10 to a far lesser degree, some of the traits associated with vocal learning also exist
11 in mice (Arriaga and Jarvis, 2013). The hypothesis presented here is grounded in
12 findings from the avian literature on song learning. There are a number of neural
13 and behavioral parallels between humans and songbirds (see Bertram et al.,
14 2014; Brainard and Doupe, 2013; Doupe and Kuhl, 1999; Fee and Goldberg,
15 2011; Mooney, 2009; Sakata and Vehrencamp, 2012; Woolley and Kao, 2015). In
16 the same way as human infants learning speech, songbirds also begin vocal
17 learning with a perception phase, during which they listen to songs from a tutor
18 (Brainard and Doupe, 2000; Doupe and Kuhl, 1999; Konishi, 2004). Without
19 exposure to adult song, production of accurate vocalizations is not possible. The
20 production phase in songbirds begins with 'subsong', (similar to human
21 babbling), before moving onto 'plastic song' (while they practice what they are
22 learning), before 'crystallized' song (the equivalent of human native speech)
23 appears. During the plastic song stage, songbirds use trial-and-error learning to
24 adjust their vocal performance until the auditory feedback from their vocal
25 output matches the auditory templates acquired during the auditory learning
26 phase (Bolhuis et al., 2010; Brainard and Doupe, 2000; Mooney, 2009).

27
28 As well as similarities in the developmental progression of learning, human
29 speech learning and birdsong acquisition have parallels at the neural and genetic
30 levels (Bolhuis et al., 2010; Jarvis et al., 2005; Ölveczky et al., 2005; Pfenning et
31 al., 2014). A recent gene expression study examined transcriptional
32 specializations in humans and song-learning birds and found that the songbird
33 RA nucleus is most similar to layer 5 neurons of human laryngeal motor cortex
34 (LMC) (Pfenning et al., 2014). The songbird Area X in the striatum is most similar
35 to a region within the human anterior striatum (Pfenning et al., 2014), and data
36 from our recent vocal learning study on humans support this finding (Simmonds
37 et al., 2014). The songbird HVC is similar to layers 2 and 3 neurons of primary
38 motor cortex, and thereby possibly also to LMC; songbird LMAN has a weak
39 similarity to Broca's area that requires further investigation for confirmation;
40 DLM (dorsolateral nucleus of the medial thalamus) is most similar to the human
41 anterior thalamus necessary for speech learning and production (Jarvis, 2004;
42 Petkov and Jarvis, 2012).

43
44 In this paper I present a hypothesis on how foreign accents could be improved
45 by optimizing variability in vocal learning brain circuits, followed by support for
46 the hypothesis, drawing on the literature on variability in songbird vocal
47 learning and variability in motor learning. The paper concludes with approaches
48 for testing the hypothesis.

1 **Hypothesis (Figure 2)**

2 The hypothesis presented here is that, as songbirds do, humans have a vocal
3 learning pathway that controls neural and behavioral variability and the
4 influence of this pathway is reduced in older learners, which leads to an inability
5 to master the native accent when learning new languages. Furthermore, if this
6 variability can be optimized in late learners, vocal learning could perhaps be
7 more complete and thereby reduce or eliminate the foreign accent. The focus
8 here is on variability in the acoustic structure of speech, rather than sequencing
9 or timing variability.

10

11 *i: The vocal learning pathway in late language learners becomes inactive*
12 *too early in the learning process and prevents accurate pronunciation in a*
13 *foreign language*

14 In 2004, Erich Jarvis (Jarvis 2004, 2007) put forward the hypothesis that as in
15 songbirds there exist two pathways in humans: one for vocal learning, and one
16 for production of previously learnt speech. Learning novel motor sequences that
17 are necessary for accurately pronouncing foreign speech is a challenge, and in
18 this paper, I argue that for late learners of a foreign language, the vocal learning
19 pathway becomes inactive too early in the learning process, engaging the motor
20 pathway once again. Consequently these late learners do not acquire novel
21 sequences of articulatory movements for the new speech; instead they adapt
22 existing production sequences, which results in speaking the new language with
23 an accent influenced by their own first language, rather than mastering the
24 native-like accent of the target language.

25

26 Figure 2A presents a simplified diagram of the motor and vocal learning
27 pathways in songbirds and humans. In both songbird pathways the HVC
28 ultimately projects to motor neurons in the brainstem (the nXllts), which then
29 projects to the vocal muscles for vocalization. Following the vocal motor
30 pathway, the HVC projects directly to the RA, which in turn makes a direct
31 projection to brainstem vocal motor neurons (see Figure 2Ai). The vocal learning
32 pathway (anterior forebrain pathway – AFP) consists of a cortical-basal-ganglia-
33 thalamic loop similar to mammals, involving Area X, the DLM and LMAN (Jarvis,
34 2004, 2006). This loop can be further segregated into lateral and medial loops,
35 both receiving input from HVC into Area X, but with different outputs. The
36 output of the lateral loop is from LMAN to RA; the output of the medial loop is
37 from MMAN (medial magnocellular nucleus of the midbrain) to HVC (Jarvis,
38 2006). The HVC continues developing until month four post-hatch, near the end
39 of the plastic-song stage (Alvarez-Buylla et al., 1992).

40

41 In songbirds the vocal learning pathway is involved during the acquisition of the
42 song pattern and remains important for the modulation of song across social
43 contexts. The vocal motor pathway is involved in producing the learned song
44 (Nottebohm, 2005), and during the plastic song stage in juveniles both pathways
45 interact (Ölveczky et al., 2005). Subsong in juvenile birds does not require HVC, a
46 key premotor area for singing in adult birds, but does require activity in RA and
47 LMAN, which is involved in learning but is not necessary for adult singing of an
48 established song (Aronov et al., 2008). Therefore the relative contributions of the
49 vocal motor and learning pathways seem to change across development in

1 songbirds. It is likely that a similar shift in balance between the two pathways
2 occurs in humans at different stages of learning. I suggest that in late learners of
3 a foreign language, the vocal learning pathway is involved to a greater extent at
4 the beginning of the learning phase but before learning is complete, the balance
5 in activity between the two circuits shifts more to the motor pathway once again,
6 which prevents accurate learning of pronunciation.

7
8 'Closed-ended learners', such as the zebra finch, are unable to learn a new song
9 in adulthood, even with an intact AFP (Brainard and Doupe, 2002; Funabiki and
10 Funabiki, 2009), as the song they learn becomes crystallized at around 90 days
11 post-hatch and remains stable throughout adulthood (Brainard, 2002). An
12 'open-ended learner', such as a canary, is able to repeat the learning process in
13 adulthood (Brainard, 2002; Nottebohm et al., 1976). If a region within the vocal
14 learning pathway is lesioned in an adult open-ended learner, the bird can
15 continue to produce song it had previously learnt, but is unable to learn a new
16 song (Brainard, 2004; Brainard and Doupe, 2000, 2002). In humans, subcortical
17 structures including the basal ganglia, similar to regions within the songbird
18 AFP, modulate production of overlearned language (e.g. poems or quotations),
19 automatic speech (e.g. counting or reciting the days of the week) and formulaic
20 expressions or fillers (Bridges et al., 2013). Patients with lesions in these regions
21 produce fewer examples of formulaic language than controls (Sidtis et al., 2009).
22 This suggests that overlearned language relies more on subcortical structures
23 than novel language does, perhaps reflecting less reliance on the vocal learning
24 pathway in later language learning.

25
26 *ii: Variability in neural activity within vocal motor circuitry generates*
27 *vocal variability that supports the acquisition of native-like pronunciation*
28 *in a foreign language*

29 Further, I suggest that prolonged random variation is an essential prerequisite
30 for vocal learning, and optimal variability within the vocal learning pathway
31 generates vocal variability and supports accurate pronunciation with a native-
32 like accent. Activity within the vocal learning pathway in adult songbirds
33 remains important for real-time generation of spectral variability necessary for
34 adapting the song based on different social contexts. In songbirds, vocal
35 variability is actively injected into the premotor song-control region RA (robust
36 nucleus of the arcopallium) by the LMAN (lateral magnocellular nucleus of the
37 anterior nidopallium), which is the output of the vocal learning pathway
38 (Goldberg and Fee, 2011). The LMAN is not necessary for the production of song,
39 only learning and modification to it. When LMAN neurons are inactive, the vocal
40 motor pathway produces an accurate, established pattern. When the LMAN is
41 active during song production, there is much more variability in the song. This
42 variability is needed to reach accurate imitation of a pattern. I argue that in
43 humans, strategies that increase the variability of neural activity in the vocal
44 learning pathway may increase behavioral variability and exploration and
45 promote more successful learning.

46
47 Figure 2B presents suggested levels of vocal variability when using the two
48 pathways. I suggest that when using the motor pathway, production is stable,
49 with little trial-by-trial variability. When using the vocal learning pathway, trial-

1 by-trial variability gradually increases, reflecting an exploratory phase in which
2 the learners try out different ways of pronouncing the words, before decreasing
3 and stabilizing once the 'best' performance has been identified.

4 **Support for the hypothesis**

5 In this paper I argue that if variability can be optimized in late language learners,
6 vocal learning could perhaps be more complete and thereby enable mastery of a
7 native-like accent in the foreign language. It is not simply that variability in vocal
8 learning needs to increase. Too much variability, or noise, prevents learning just
9 as too little does (Faisal et al., 2008). Therefore, for effective learning it is
10 necessary to optimize the amount of variability. By trying different versions of
11 producing the target, a learner is able to monitor outcomes and refine the
12 movement sequences that result in the most desired outcome. This is true in
13 songbirds and is likely true in humans as well.

14
15 *Variability in songbird vocal learning*

16 A critical amount of noise within the song production pathway is necessary
17 during song learning (Doya and Sejnowski, 1995; Ölveczky et al., 2005) and song
18 variability is generated by the AFP (Woolley and Kao, 2015). This variability in
19 the AFP has been shown to correlate with performance variability (Kao et al.,
20 2005; Woolley and Kao, 2015). Although a critical amount of noise appears
21 essential for songbird learning, optimal learning will only occur within the
22 appropriate level of noise for a given stage of learning. There is a reduction in
23 variability within the AFP as the song crystalizes, although some neural and
24 vocal stochastic variability is present even in adult songbirds with apparently
25 stable song (Andalman and Fee, 2009; Kao and Brainard, 2006; Kao et al., 2005).
26 Using altered auditory feedback in adult songbirds, Tumer and Brainard
27 observed that birds were able to learn how their song changed as a result of
28 small variations in vocal performance (Tumer and Brainard, 2007). They suggest
29 that residual variability that persists in well-learned skills reflects motor
30 exploration as part of the trial-and-error learning and monitoring processes, and
31 that this helps to support continuous learning and optimization of performance.

32
33 Within the AFP song learning pathway, lesions to Area X have little or no effect
34 on song variability during the vocal babbling stage (Goldberg and Fee, 2011), but
35 when Area X is lesioned in juveniles, the song does not fully crystallize as they
36 become adults and instead remains variable (Scharff and Nottebohm, 1991;
37 Sohrabji et al., 1990). In contrast, LMAN inactivation results in reduced, almost
38 absent variability in song in juveniles and adults (Aronov et al., 2008; Kao and
39 Brainard, 2006; Kao et al., 2005; Ölveczky et al., 2005; Thompson et al., 2011).
40 Young birds at an early stage of song development, which have the most variable
41 song performance, show the greatest reduction in song variability following
42 LMAN inactivation (Ölveczky et al., 2005). Similarly, during vocal babbling in
43 juveniles a lesion to the DLM, part of the thalamus that receives output from the
44 basal ganglia, almost completely removes variability and causes the birds to
45 produce a stable stereotyped song (Goldberg and Fee, 2011).

46
47 A decrease in variability has also been observed following lesions to the dorsal
48 arcopallium, adjacent to RA, by authors who suggest this to be an auditory region

1 involved in song learning (Bottjer and Altenau, 2010). However, this region,
2 along with other brain areas adjacent to the vocal systems of vocal learning
3 birds, has been shown to be active during limb and body movements (Feenders
4 et al., 2008). This suggests that the systems involved in vocalizations are
5 controlled by a cerebral motor system. Although a similar auditory pathway
6 exists in both vocal learners and non-learners, vocal learners have a specialized
7 vocal motor system that enables auditory input to be translated into vocal
8 signals (Feenders et al., 2008). A recent electrophysiology and lesion study
9 supports this motor hypothesis, again showing motor behavior and movement
10 control of this region (Mandelblat-Cerf et al., 2014). Further support for the
11 motor hypothesis comes from Pfenning and colleagues who, using gene
12 expression, found that the molecular profile of this region is similar to that of the
13 motor and premotor cortex in primates, and not the auditory cortex (Pfenning et
14 al., 2014). Therefore, the variability observed by Bottjer and Altenau may be
15 similar to that found in RA and motor pathways.

16
17 In trial-and-error learning in juvenile songbirds the 'trial' is represented by the
18 variability in the song, reflecting the motor exploration phase, and the 'error' is
19 represented by evaluation of song performance, based on auditory feedback
20 (Andalman and Fee, 2009; Fee and Goldberg, 2011; Sober and Brainard, 2009;
21 Tumer and Brainard, 2007). Such variability is necessary for reinforcement-
22 based trial-and-error learning, as the learning process requires exploration of a
23 range of action sequences, evaluation of performance with each and
24 modifications to behavior that result in improved performance (Ölveczky et al.,
25 2005).

26
27 Even in crystallized song in adult birds, trial-by-trial variability persists. This
28 variability supports ongoing motor exploration, which maintains performance
29 and makes modifications when necessary (Tumer and Brainard, 2007). Song
30 variability is also context-dependent. During 'directed' song, in which a male
31 sings a courtship song to a female, the sequencing and structure of syllables are
32 much less variable than when the male sings alone ('undirected' song) (Kao and
33 Brainard, 2006; Kao et al., 2005; Kojima and Doupe, 2011; Ölveczky et al., 2005;
34 Sakata et al., 2008; Teramitsu and White, 2006; Woolley et al., 2014). This
35 suggests that singing alone reflects a practice state of exploratory vocal learning,
36 and directed singing reflects a performance state, in which the male produces the
37 best rendition of their song they memorized during the sensitive period in
38 development (Kojima and Doupe, 2011). LMAN activity is much greater and
39 more variable during undirected song than during directed song (Brainard and
40 Doupe, 2013; Hessler and Doupe, 1999; Kao and Brainard, 2006; Kao et al., 2005;
41 Kojima and Doupe, 2011; Woolley and Kao, 2015) and a lesion to LMAN removes
42 the variability and causes undirected singing to be much more consistent
43 (Hampton et al., 2009; Kao and Brainard, 2006; Kao et al., 2005).

44 45 *Variability in motor learning*

46 The hypothesis proposed here is also supported by findings from research on
47 motor learning more generally. Noise in general motor learning (not just vocal
48 learning) has been defined as a mismatch between expected and actual sensory
49 feedback that is not necessarily related to performance errors (Faisal et al.,

1 2008). Recent theoretical and experimental work suggests an important role for
2 noise, termed stochastic facilitation, in motor learning, i.e., variability or noise in
3 the motor movement is necessary for the development of expertise (McDonnell
4 and Ward, 2011; Mendez-Balbuena et al., 2012). Stochastic processes,
5 introducing variability in the execution of motor movements, permit a full
6 exploration of the learning space. Motor learning involves an 'exploration' phase,
7 during which trial-and-error learning is performed to identify the optimal
8 movement for a successful outcome. Once that is identified, the learner moves
9 into the 'exploitation' phase, in which they continue producing that movement
10 until the necessary outcome is achieved. Motor learning therefore involves a
11 tradeoff between performing multiple movements to find the one that most
12 reliably produces the desired outcome, and continuing to produce that
13 movement once it has been identified (Müller and Sternad, 2009; Ravbar et al.,
14 2012). During the exploration phase performance is highly variable, and it
15 becomes more consistent when the average performance is closer to the target
16 outcome, suggesting that variance decreases with the bias (Müller and Sternad,
17 2009; Ravbar et al., 2012). The tradeoff between exploration and stabilization is
18 not the same throughout the learning process. When learning continuous actions
19 (such as dancing), different components of the action may need exploratory
20 variability while others, which may be closer to the target, require stabilization
21 (Doya, 2000). With this type of approach, breaking the movements down into
22 segments would allow variability to be regulated locally so that only those parts
23 of the action that need to change the most undergo exploration, i.e. learning
24 based on the local bias (Doya, 2000; Ravbar et al., 2012).

25
26 Individual differences in the amount of motor variability have been associated
27 with the ability to learn or adapt motor skills (Sober and Brainard, 2012; Wu et
28 al., 2014) and models of trial-and-error learning suggest that previous
29 performance can predict the amount of variability in the motor output (Kao et
30 al., 2008). This suggests that motor 'noise', or variability, is a central component
31 of motor learning (Herzfeld and Shadmehr, 2014). Neural variability is also an
32 indicator of motor learning. As motor habits form, spike firing in the
33 ventromedial striatum peaks at the beginning and end of the motor sequence,
34 and changes to this firing have been suggested to be a sign of learning (Howe et
35 al., 2011). In non-vocal motor learning in rodents, using a reward-based
36 conditional T-maze task, spiking of striatal neurons has been shown to be highly
37 variable at the initial stage of learning, but following training became more
38 consistent (Barnes et al., 2005). The variable firing rate during learning is
39 considered to represent 'neural exploration', whereas the stable firing after
40 learning reflects 'neural exploitation'.

41 42 **Testing the hypothesis**

43 This converging literature from research on songbird vocal learning and more
44 general motor learning motivated our previous work, which suggests that in late
45 learners of a second language, the vocal learning pathway may become inactive
46 too early, ending the motor learning phase prematurely. Instead, the motor
47 pathway is recruited once more, which results in the learner producing the
48 original native motor patterns for speech; this results in speaking with a foreign
49 accent (Simmonds et al., 2014). The hypothesis proposed here could be tested

1 using behavioral interventions that keep speakers in the learning phase
2 (engaging the vocal learning pathway) for longer, with the prediction that this
3 would allow them to develop new motor patterns that result in more native-like
4 accuracy of pronunciation. This could be investigated using strategies that
5 induce neural and behavioral variability, such as altering the auditory feedback
6 that learners receive. Disrupting auditory feedback in songbirds results in rapid
7 changes to learned song (Andalman and Fee, 2009; Hoffmann and Sober, 2014;
8 Tumer and Brainard, 2007), although variability itself did not increase in these
9 studies. This suggests that altering auditory feedback induces experimentally
10 controlled 'errors' and changes in song performance (Andalman and Fee, 2009;
11 Fee and Goldberg, 2011; Tumer and Brainard, 2007). Sakata and Brainard have
12 also found populations of neurons that appear to be sensitive to auditory
13 feedback (Sakata and Brainard, 2008). Dramatic changes to auditory feedback
14 can increase song variability and decrystallize the song (Leonardo and Konishi,
15 1999). Using Bengalese finches, Woolley and Rubel have demonstrated that
16 temporary deafening leads to the rapid deterioration of syllable structure and an
17 increase in vocal variability, but once hearing is restored, song is produced
18 normally again (Woolley and Rubel, 1997, 2002). Therefore, although altered
19 auditory feedback disrupts speech production, the auditory template of the
20 acoustic template could remain intact. Assessing speech perception as well as
21 production would identify whether the motor pattern or auditory target has
22 been impaired.

23
24 Altered auditory feedback has also been shown to affect vocal production in
25 humans (Houde and Jordan, 1998; Jones and Munhall, 2005; Kort et al., 2014;
26 Lametti et al., 2012; Ogane and Honda, 2014; Tourville et al., 2008), although its
27 role in language learning has not been explored. Different types of feedback
28 could be used to investigate different ways of modulating variability during vocal
29 learning, manipulating cognitive and motor processes to promote variability.
30 Types of auditory feedback could include frequency-altered, delayed,
31 background noise or white noise. Behavioral variability could be assessed by
32 analyzing the acoustic properties of participants' speech, including simple
33 measures of intensity, duration and frequency, as well as correlations of the
34 long-term spectra of specific words and characterization of formants.
35 Somatosensory feedback could also be manipulated, for example altering jaw
36 movements during speech, which has been shown to result in a mismatch
37 between the expected sensations and the sensory feedback actually received,
38 which causes somatosensory error signals that lead to compensatory
39 movements (Guenther et al., 2006; Tourville et al., 2005). Some speakers rely on
40 auditory feedback information and others rely more on somatosensory feedback
41 (Lametti et al., 2012). Investigating a range of alterations to feedback would
42 allow optimization of variability.

43
44 Using continuous speech at the sentence level would allow evaluation of
45 performance to be carried out locally, focusing on specific words or phonemes.
46 Rather than aiming to adapt a speaker's overall level of variability, altered
47 feedback could be used to only induce motor exploration in sounds that need to
48 change. By assessing an individual's speech, feedback manipulations could be
49 developed to only occur for certain words. This type of approach has previously

1 been investigated in zebra finches by manipulating song learning so that only a
2 specific part of the song requires vocal exploration. Ravbar and colleagues found
3 no apparent increase in the variability of one syllable when a second first
4 appeared, demonstrating that the bird was able to rapidly switch between
5 performing a highly stereotyped and a highly variable syllable (Ravbar et al.,
6 2012).

7
8 The hypothesis proposed here could also be tested using neurobiologically-
9 plausible computational simulations of the neural systems involved in vocal
10 learning. The known neuroanatomy and structural connections of networks
11 involved in speech production, defined using imaging studies, could be used to
12 create a neuroanatomically-constrained model to simulate behavioral variability
13 and learning effects. This type of model would help explain how neural and
14 behavioral stochastic facilitation, with a focus on the striatum as a mediator,
15 could affect vocal learning and allow us to explore, theoretically, the most
16 effective amount of stochastic variability for successful learning. This would also
17 allow for theoretical investigation of the influence of stochastic processes on
18 learning and to simulate interventions in order to predict the optimal level of
19 induced variability for best learning. Larger projects could then investigate the
20 long-term benefits of these novel strategies for foreign language learning, which
21 could lead to the development of new training materials with a strong evidence
22 base, and discussions with educational policy-makers directing future strategies
23 for improving foreign language learning outcomes.

24 25 26 27 **References**

28
29 Alario, F.X., Goslin, J., Michel, V., and Laganaro, M. (2010). The functional origin of
30 the foreign accent. *Psychol Sci* 21, 15-20.

31 Alvarez-Buylla, A., Ling, C.-Y., and Nottebohm, F. (1992). High vocal center
32 growth and its relation to neurogenesis, neuronal replacement and song
33 acquisition in juvenile canaries. *J Neurobiol* 23, 396-406.

34 Andalman, A.S., and Fee, M.S. (2009). A basal ganglia-forebrain circuit in the
35 songbird biases motor output to avoid vocal errors. *Proc Natl Acad Sci U S A* 106,
36 12518-12523.

37 Aronov, D., Andalman, A.S., and Fee, M.S. (2008). A Specialized Forebrain Circuit
38 for Vocal Babbling in the Juvenile Songbird. *Science* 320, 630-634.

39 Arriaga, G., and Jarvis, E.D. (2013). Mouse vocal communication system: Are
40 ultrasounds learned or innate? *Brain Lang* 124, 96-116.

41 Barnes, T.D., Kubota, Y., Hu, D., Jin, D.Z., and Graybiel, A.M. (2005). Activity of
42 striatal neurons reflects dynamic encoding and recoding of procedural
43 memories. *Nature* 437, 1158-1161.

- 1 Bertram, R., Daou, A., Hyson, R.L., Johnson, F., and Wu, W. (2014). Two neural
2 streams, one voice: Pathways for theme and variation in the songbird brain.
3 *Neuroscience* 277, 806-817.
- 4 Bley-Vroman, R. (1990). The logical problem of foreign language learning.
5 *Linguistic Analysis* 20, 3-49.
- 6 Bolhuis, J.J., Okanoya, K., and Scharff, C. (2010). Twitter evolution: converging
7 mechanisms in birdsong and human speech. *Nat Rev Neurosci* 11, 747-759.
- 8 Bottjer, S.W., and Altenau, B. (2010). Parallel pathways for vocal learning in basal
9 ganglia of songbirds. *Nat Neurosci* 13, 153-155.
- 10 Brainard, M.S. (2004). Contributions of the Anterior Forebrain Pathway to Vocal
11 Plasticity. *Ann N Y Acad Sci* 1016, 377-394.
- 12 Brainard, M.S., and Doupe, A.J. (2000). Interruption of a basal ganglia-forebrain
13 circuit prevents plasticity of learned vocalizations. *Nature* 404, 762-766.
- 14 Brainard, M.S., and Doupe, A.J. (2002). What songbirds teach us about learning.
15 *Nature* 417, 351-358.
- 16 Brainard, M.S., and Doupe, A.J. (2013). Translating Birdsong: Songbirds as a
17 model for basic and applied medical research. *Annu Rev Neurosci* 36, 489-517.
- 18 Bridges, K.A., Van Lancker Sidtis, D., and Sidtis, J.J. (2013). The role of subcortical
19 structures in recited speech: Studies in Parkinson's disease. *J Neurolinguist* 26,
20 591-601.
- 21 Doupe, A.J., and Kuhl, P.K. (1999). Birdsong and Human Speech: Common
22 Themes and Mechanisms. *Annu Rev Neurosci* 22, 567-631.
- 23 Doya, K. (2000). Complementary roles of basal ganglia and cerebellum in
24 learning and motor control. *Curr Opin Neurobiol* 10, 732-739.
- 25 Doya, K., and Sejnowski, T.J. (1995). A novel reinforcement model of birdsong
26 vocalization learning. In *Advances in neural information processing systems* 7, G.
27 Tesauro, D.S. Touretzky, and T.K. Leen, eds. (Massachusetts Institute of
28 Technology).
- 29 Faisal, A.A., Selen, L.P.J., and Wolpert, D.M. (2008). Noise in the nervous system.
30 *Nat Rev Neurosci* 9, 292-303.
- 31 Fee, M.S., and Goldberg, J.H. (2011). A hypothesis for basal ganglia-dependent
32 reinforcement learning in the songbird. *Neuroscience* 198, 152-170.

- 1 Feenders, G., Liedvogel, M., Rivas, M., Zapka, M., Horita, H., Hara, E., Wada, K.,
2 Mouritsen, H., and Jarvis, E.D. (2008). Molecular Mapping of Movement-
3 Associated Areas in the Avian Brain: A Motor Theory for Vocal Learning Origin.
4 PLoS ONE 3, e1768.
- 5 Flege, J.E. (1995). Second language speech learning: theory, findings, and
6 problems. In Speech perception and linguistic experience: issues in cross-
7 language research, W. Strange, ed. (Timonium, MD, York Press), pp. 233-277.
- 8 Funabiki, Y., and Funabiki, K. (2009). Factors limiting song acquisition in adult
9 zebra finches. *Developmental Neurobiology* 69, 752-759.
- 10 Goldberg, J.H., and Fee, M.S. (2011). Vocal babbling in songbirds requires the
11 basal ganglia-recipient motor thalamus but not the basal ganglia. *J Neurophysiol*
12 105, 2729-2739.
- 13 Golestani, N., Molko, N., Dehaene, S., LeBihan, D., and Pallier, C. (2007). Brain
14 structure predicts the learning of foreign speech sounds. *Cereb Cortex* 17, 575-
15 582.
- 16 Golestani, N., and Pallier, C. (2006). Anatomical Correlates of Foreign Speech
17 Sound Production. *Cereb Cortex* 17, 929-934.
- 18 Graybiel, A.M. (2008). Habits, Rituals, and the Evaluative Brain. *Annu Rev*
19 *Neurosci* 31, 359-387.
- 20 Guenther, F.H., Ghosh, S.S., and Tourville, J.A. (2006). Neural modeling and
21 imaging of the cortical interactions underlying syllable production. *Brain Lang*
22 96, 280-301.
- 23 Hampton, C.M., Sakata, J.T., and Brainard, M.S. (2009). An Avian Basal Ganglia-
24 Forebrain Circuit Contributes Differentially to Syllable Versus Sequence
25 Variability of Adult Bengalese Finch Song. *J Neurophysiol* 101, 3235-3245.
- 26 Herzfeld, D.J., and Shadmehr, R. (2014). Motor variability is not noise, but grist
27 for the learning mill. *Nat Neurosci* 17, 149-150.
- 28 Hessler, N.A., and Doupe, A.J. (1999). Social context modulates singing-related
29 neural activity in the songbird forebrain. *Nat Neurosci* 2, 209-211.
- 30 Hikosaka, O., Nakahara, H., Rand, M.K., Sakai, K., Lu, X., Nakamura, K., Miyachi, S.,
31 and Doya, K. (1999). Parallel neural networks for learning sequential
32 procedures. *Trends Neurosci* 22, 464-471.
- 33 Hikosaka, O., Nakamura, K., Sakai, K., and Nakahara, H. (2002). Central
34 mechanisms of motor skill learning. *Curr Opin Neurobiol* 12, 217-222.

- 1 Hoffmann, L.A., and Sober, S.J. (2014). Vocal Generalization Depends on Gesture
2 Identity and Sequence. *J Neurosci* 34, 5564-5574.
- 3 Houde, J.F., and Jordan, M.I. (1998). Sensorimotor adaptation in speech
4 production. *Science* 279, 1213-1216.
- 5 Howe, M.W., Atallah, H.E., McCool, A., Gibson, D.J., and Graybiel, A.M. (2011).
6 Habit learning is associated with major shifts in frequencies of oscillatory
7 activity and synchronized spike firing in striatum. *Proc Natl Acad Sci U S A* 108,
8 16801-16806.
- 9 Jarvis, E. (2007). Neural systems for vocal learning in birds and humans: a
10 synopsis. *J Ornithol* 148, 35-44.
- 11 Jarvis, E.D. (2004). Learned birdsong and the neurobiology of human language.
12 *Ann N Y Acad Sci* 1016, 749-777.
- 13 Jarvis, E.D. (2006). Selection for and against vocal learning in birds and
14 mammals. *Ornithol Sci* 5, 5-14.
- 15 Jarvis, E.D., Gunturkun, O., Bruce, L., Csillag, A., Karten, H., Kuenzel, W., Medina, L.,
16 Paxinos, G., Perkel, D.J., Shimizu, T., *et al.* (2005). Avian brains and a new
17 understanding of vertebrate brain evolution. *Nat Rev Neurosci* 6, 151-159.
- 18 Jarvis, E.D., Ribeiro, S., da Silva, M.L., Ventura, D., Vielliard, J., and Mello, C.V.
19 (2000). Behaviourally driven gene expression reveals song nuclei in
20 hummingbird brain. *Nature* 406, 628-632.
- 21 Jones, J.A., and Munhall, K.G. (2005). Remapping auditory-motor representations
22 in voice production. *Curr Biol* 15, 1768-1772.
- 23 Jueptner, M., and Weiller, C. (1998). A review of differences between basal
24 ganglia and cerebellar control of movements as revealed by functional imaging
25 studies. *Brain* 121, 1437-1449.
- 26 Kao, M.H., and Brainard, M.S. (2006). Lesions of an Avian Basal Ganglia Circuit
27 Prevent Context-Dependent Changes to Song Variability. *J Neurophysiol* 96,
28 1441-1455.
- 29 Kao, M.H., Doupe, A.J., and Brainard, M.S. (2005). Contributions of an avian basal
30 ganglia-forebrain circuit to real-time modulation of song. *Nature* 433, 638-643.
- 31 Kao, M.H., Wright, B.D., and Doupe, A.J. (2008). Neurons in a Forebrain Nucleus
32 Required for Vocal Plasticity Rapidly Switch between Precise Firing and Variable
33 Bursting Depending on Social Context. *J Neurosci* 28, 13232-13247.

- 1 Kojima, S., and Doupe, A.J. (2011). Social performance reveals unexpected vocal
2 competency in young songbirds. *Proc Natl Acad Sci U S A* *108*, 1687-1692.
- 3 Konishi, M. (2004). The Role of Auditory Feedback in Birdsong. *Ann N Y Acad Sci*
4 *1016*, 463-475.
- 5 Kort, N.S., Nagarajan, S.S., and Houde, J.F. (2014). A bilateral cortical network
6 responds to pitch perturbations in speech feedback. *Neuroimage* *86*, 525-535.
- 7 Kuhl, P.K. (2004). Early language acquisition: cracking the speech code. *Nat Rev*
8 *Neurosci* *5*, 831-843.
- 9 Lametti, D.R., Nasir, S.M., and Ostry, D.J. (2012). Sensory preference in speech
10 production revealed by simultaneous alteration of auditory and somatosensory
11 feedback. *J Neurosci* *32*, 9351-9358.
- 12 Leonardo, A., and Konishi, M. (1999). Decrystallization of adult birdsong by
13 perturbation of auditory feedback. *Nature* *399*, 466-470.
- 14 Mandelblat-Cerf, Y., Las, L., Denisenko, N., and Fee, M. (2014). A role for
15 descending auditory cortical projections in songbird vocal learning. *eLIFE*
16 *10.7554/eLife.02152*.
- 17 McDonnell, M.D., and Ward, L.M. (2011). The benefits of noise in neural systems:
18 bridging theory and experiment. *Nat Rev Neurosci* *12*, 415-426.
- 19 Mendez-Balbuena, I., Manjarrez, E., Schulte-Mönting, Huethe, F., Tapia, J.A., Hepp-
20 Reymond, M.-C., and Kristeva, R. (2012). Improved sensorimotor performance
21 via stochastic resonance. *J Neurosci* *32*, 12612-12618.
- 22 Miyachi, S., Hikosaka, O., Miyashita, K., Kárádi, Z., and Rand, M.K. (1997).
23 Differential roles of monkey striatum in learning of sequential hand movement.
24 *Exp Brain Res* *115*, 1-5.
- 25 Mooney, R. (2009). Neural mechanisms for learned birdsong. *Learn Mem* *16*,
26 655-669.
- 27 Müller, H., and Sternad, D. (2009). Motor Learning: Changes in the Structure of
28 Variability in a Redundant Task. In *Progress in Motor Control*, D. Sternad, ed.
29 (Springer US), pp. 439-456.
- 30 Nottebohm, F. (2005). The neural basis of birdsong. *PLoS Biol* *3*, e164.
- 31 Nottebohm, F., Stokes, T.M., and Leonard, C.M. (1976). Central control of song in
32 the canary, *Serinus canarius*. *J Comp Neurol* *165*, 457-486.

- 1 Ogane, R., and Honda, M. (2014). Speech Compensation for Time-Scale-Modified
2 Auditory Feedback. *J Speech Lang Hear Res* 57, S616-S625.
- 3 Ölveczky, B.P., Andalman, A.S., and Fee, M.S. (2005). Vocal experimentation in the
4 juvenile songbird requires a basal ganglia circuit. *PLoS Biol* 3, e153.
- 5 Petkov, C.I., and Jarvis, E. (2012). Birds, Primates and Spoken Language Origins:
6 Behavioral Phenotypes and Neurobiological Substrates. *Front Evol Neurosci*
7 4:12. doi: 10.3389/fnevo.2012.00012.
- 8 Pfenning, A.R., Hara, E., Whitney, O., Rivas, M.V., Wang, R., Roulhac, P.L., Howard,
9 J.T., Wirthlin, M., Lovell, P.V., Ganapathy, G., *et al.* (2014). Convergent
10 transcriptional specializations in the brains of humans and song-learning birds.
11 *Science* 346.
- 12 Piske, T., MacKay, I.R.A., and Flege, J.E. (2001). Factors affecting degree of foreign
13 accent in an L2: a review. *J Phon* 29, 191-215.
- 14 Ravbar, P., Lipkind, D., Parra, L.C., and Tchernichovski, O. (2012). Vocal
15 Exploration Is Locally Regulated during Song Learning. *J Neurosci* 32, 3422-
16 3432.
- 17 Sakata, J.T., and Brainard, M.S. (2008). Online contributions of auditory feedback
18 to neural activity in avian song control circuitry. *J Neurosci* 28, 11378-11390.
- 19 Sakata, J.T., Hampton, C.M., and Brainard, M.S. (2008). Social Modulation of
20 Sequence and Syllable Variability in Adult Birdsong. *J Neurophysiol* 99, 1700-
21 1711.
- 22 Sakata, J.T., and Vehrencamp, S.L. (2012). Integrating perspectives on vocal
23 performance and consistency. *J Exp Biol* 215, 201-209.
- 24 Scharff, C., and Nottebohm, F. (1991). A comparative study of the behavioral
25 deficits following lesions of various parts of the zebra finch song system:
26 implications for vocal learning. *J Neurosci* 11, 2896-2913.
- 27 Scovel, T. (1988). *A time to speak: a psycholinguistic inquiry into the critical*
28 *period for human speech* (New York, Newbury House Publishers).
- 29 Sidtis, D., Canterucci, G., and Katsnelson, D. (2009). Effects of neurological
30 damage on production of formulaic language. *Clin Linguist Phonet* 23, 270-284.
- 31 Simmonds, A.J., Leech, R., Iverson, P., and Wise, R.J.S. (2014). The response of the
32 anterior striatum during adult human vocal learning. *J Neurophysiol* 112, 792-
33 801.

- 1 Simmonds, A.J., Wise, R.J.S., Dhanjal, N.S., and Leech, R. (2011a). A comparison of
2 sensory-motor activity during speech in first and second languages. *J*
3 *Neurophysiol* 106, 470-478.
- 4 Simmonds, A.J., Wise, R.J.S., and Leech, R. (2011b). Two tongues, one brain:
5 imaging bilingual speech production. *Front Psychol* 2: 10.3389/fpsyg.2011.00166.
- 6 Sober, S.J., and Brainard, M.S. (2009). Adult birdsong is actively maintained by
7 error correction. *Nat Neurosci* 12, 927-931.
- 8 Sober, S.J., and Brainard, M.S. (2012). Vocal learning is constrained by the
9 statistics of sensorimotor experience. *Proc Natl Acad Sci U S A* 109, 21099-
10 21103.
- 11 Sohrabji, F., Nordeen, E.J., and Nordeen, K.W. (1990). Selective impairment of
12 song learning following lesions of a forebrain nucleus in the juvenile zebra finch.
13 *Behav Neural Biol* 53, 51-63.
- 14 Teramitsu, I., and White, S.A. (2006). FoxP2 Regulation during Undirected
15 Singing in Adult Songbirds. *J Neurosci* 26, 7390-7394.
- 16 Thierry, G., and Wu, Y.J. (2007). Brain potentials reveal unconscious translation
17 during foreign-language comprehension. *Proc Natl Acad Sci U S A* 104, 12530-
18 12535.
- 19 Thompson, J.A., Basista, M.J., Wu, W., Bertram, R., and Johnson, F. (2011). Dual
20 Pre-Motor Contribution to Songbird Syllable Variation. *J Neurosci* 31, 322-330.
- 21 Tourville, J.A., Guenther, F.H., Ghosh, S.S., Reilly, K.J., Bohland, J.W., and Nieto-
22 Castanon, A. (2005). Effects of acoustic and articulatory perturbation on cortical
23 activity during speech production. *NeuroImage* 26(S1), S49.
- 24 Tourville, J.A., Reilly, K.J., and Guenther, F.H. (2008). Neural mechanisms
25 underlying auditory feedback control of speech. *Neuroimage* 39, 1429-1443.
- 26 Tumer, E.C., and Brainard, M.S. (2007). Performance variability enables adaptive
27 plasticity of 'crystallized' adult birdsong. *Nature* 450, 1240-1244.
- 28 Woolley, S.C., and Kao, M.H. (2015). Variability in action: Contributions of a
29 songbird cortical-basal ganglia circuit to vocal motor learning and control.
30 *Neuroscience* 296, 39-47.
- 31 Woolley, S.C., Rajan, R., Joshua, M., and Doupe, A.J. (2014). Emergence of context-
32 dependent variability across a basal ganglia network. *Neuron* 82, 208-223.

1 Woolley, S.M.N., and Rubel, E.W. (1997). Bengalese Finches *Lonchura Striata*
2 *Domestica* Depend upon Auditory Feedback for the Maintenance of Adult Song. *J*
3 *Neurosci* 17, 6380-6390.

4 Woolley, S.M.N., and Rubel, E.W. (2002). Vocal Memory and Learning in Adult
5 Bengalese Finches with Regenerated Hair Cells. *J Neurosci* 22, 7774-7787.

6 Wu, H.G., Miyamoto, Y.R., Castro, L.N.G., Ölveczky, B.P., and Smith, M.A. (2014).
7 Temporal structure of motor variability is dynamically regulated and predicts
8 motor learning ability. *Nat Neurosci* 17, 312-321.

9 Wu, Y.J., and Thierry, G. (2010). Chinese-English bilinguals reading English hear
10 Chinese. *J Neurosci* 30, 7646-7651.

11 Wu, Y.J., and Thierry, G. (2011). Event-related brain potential investigation of
12 preparation for speech production in late bilinguals. *Front Psychol* 2, *doi*
13 *10.3389/fpsyg.2011.00114*.
14

15 Yin, H., Mulcare, S., Hilário, M., Clouse, E., Holloway, T., Davis, M., Hansson, A.,
16 Lovinger, D., and Costa, R. (2009). Dynamic reorganization of striatal circuits
17 during the acquisition and consolidation of a skill. *Nat Neurosci* 12, 333-341.
18
19
20
21

Provisional

1 **Figure Legends**

2

3 **Figure 1**

4 Figure as originally published in Petkov, C.I., and Jarvis, E. (2012), reproduced
5 with permission.

6

7 Direct and indirect vocalization pathways in complex-vocal learners, limited-
8 vocal learners and vocal non-learners

9

10 Schematic of a songbird brain **(A)** and a human brain **(B)** showing the vocal
11 motor pathway (blue arrow), the vocal learning pathway (white) and the
12 laryngeal motorneurons (red). Also shown in **(B)** is the limbic vocal pathway for
13 producing innate vocalizations (black). **(C)** Schematic of a vocal non-learning
14 bird revealing the absence of forebrain song nuclei. **(D)** Schematic of limited-
15 vocal learning monkeys showing presence of forebrain regions for innate
16 vocalization and also of an indirect projection from a ventral premotor area
17 (Area 6vr) to laryngeal motorneurons. Abbreviations: ACC, anterior cingulate
18 cortex; Am, nucleus ambiguus; Amyg, amygdala; AT, anterior thalamus; Av,
19 nucleus avalanche; DLM, dorsolateral nucleus of the medial thalamus; DM, dorsal
20 medial nucleus of the midbrain; HVC, high vocal center; LMAN, lateral
21 magnocellular nucleus of the anterior nidopallium; LMC, laryngeal motor cortex;
22 OFC, orbito-frontal cortex; PAG, periaqueductal gray; RA, robust nucleus of the of
23 arcopallium; RF, reticular formation; vPFC, ventral prefrontal cortex; VLT,
24 ventro-lateral division of thalamus; XIIIts, bird twelfth nerve nucleus.

25

26 **Figure 2**

27 Motor and vocal learning pathways in songbirds and humans and the role of
28 variability

29

30 Figure **2A**: A simplified diagram of the pathways involved in vocal learning and
31 production in songbirds and humans. **i)** In songbirds, the vocal motor pathway
32 (PDP: posterior descending pathway), used for production of established song
33 (shown in red: HVC (high vocal center), RA (robust nucleus of the arcopallium)
34 and brainstem nucleus for vocal output) is used to produce the song. The vocal
35 learning pathway (AFP: anterior forebrain pathway, shown in blue: Area X,
36 dorsolateral nucleus of the medial thalamus (DLM) and LMAN (lateral
37 magnocellular nucleus of the anterior nidopallium)) is used in songbirds during
38 the acquisition of the pattern in song learning. **ii)** In humans, the motor pathway
39 (shown in red: laryngeal motor cortex and brainstem nucleus for vocal output),
40 and the vocal learning pathway (shown in blue: anterior striatum, thalamus and
41 premotor cortex).

42 Figure **2B**: Suggested levels of vocal variability when using the two pathways. I
43 suggest that when using the motor pathway **(i)**, production is stable, with little
44 trial-by-trial variability. When using the vocal learning pathway **(ii)**, trial-by-trial
45 variability gradually increases, reflecting an exploratory phase in which the
46 learners try out different ways of pronouncing the words ('motor exploration'),

- 1 before decreasing and stabilizing once the 'best' performance has been identified
- 2 ('motor exploitation').
- 3
- 4
- 5

Provisional

1 Figure 1
2

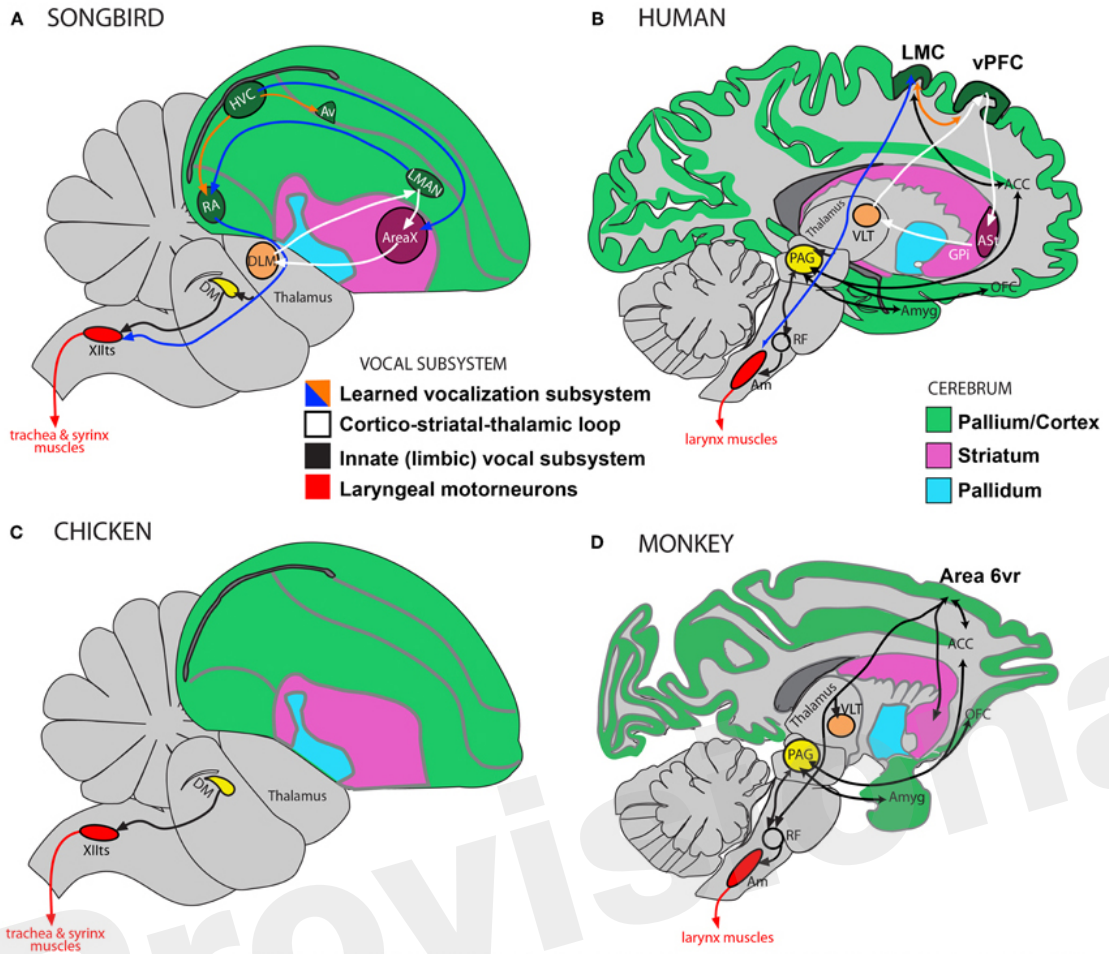


Figure as originally published in Petkov and Jarvis (2012) Front Evol Neurosci 4:12. doi: 10.3389/fnevo.2012.00012

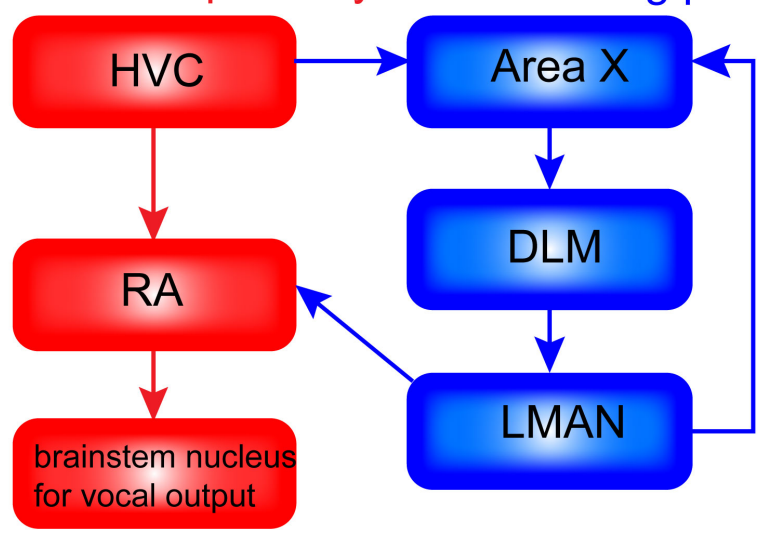
3
4

1 Figure 2

A: Vocal pathways

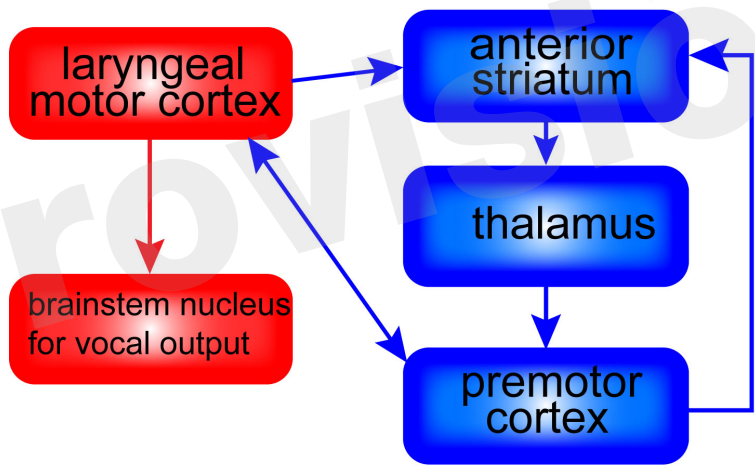
i: Songbirds

vocal motor pathway vocal learning pathway



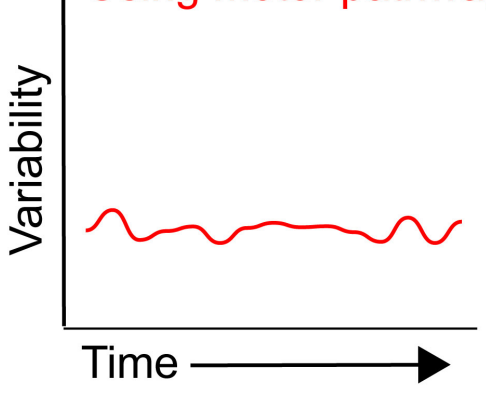
ii: Humans

vocal motor pathway vocal learning pathway

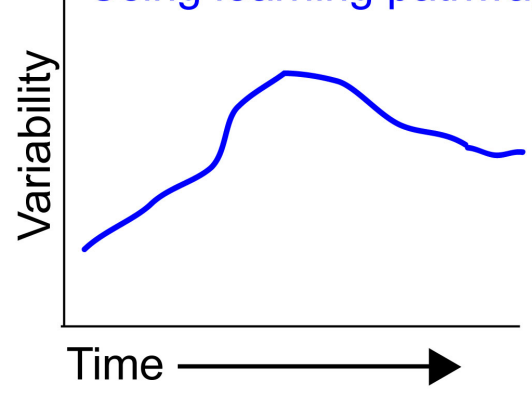


B: Vocal variability

i: Using motor pathway

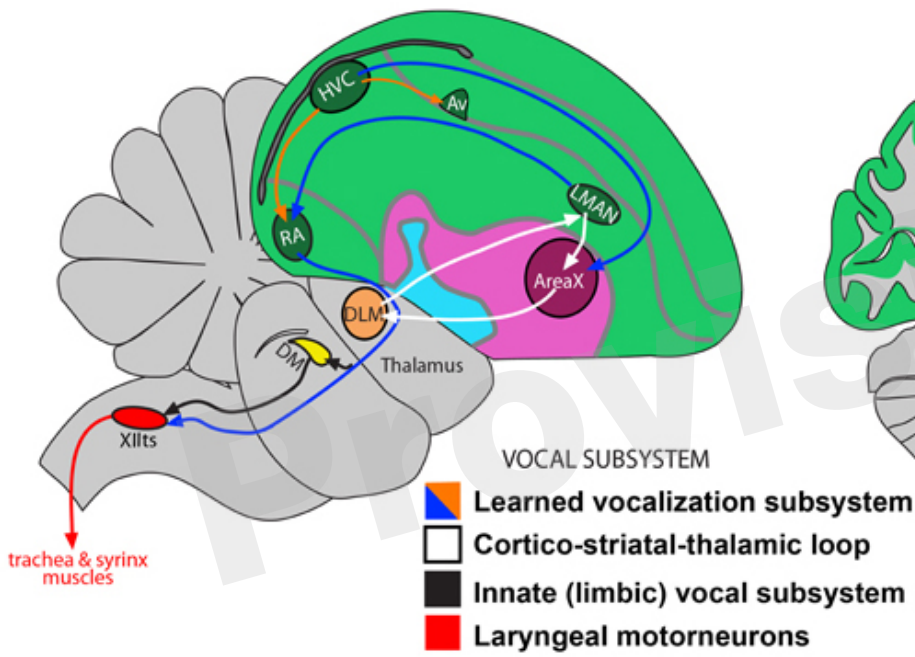


ii: Using learning pathway

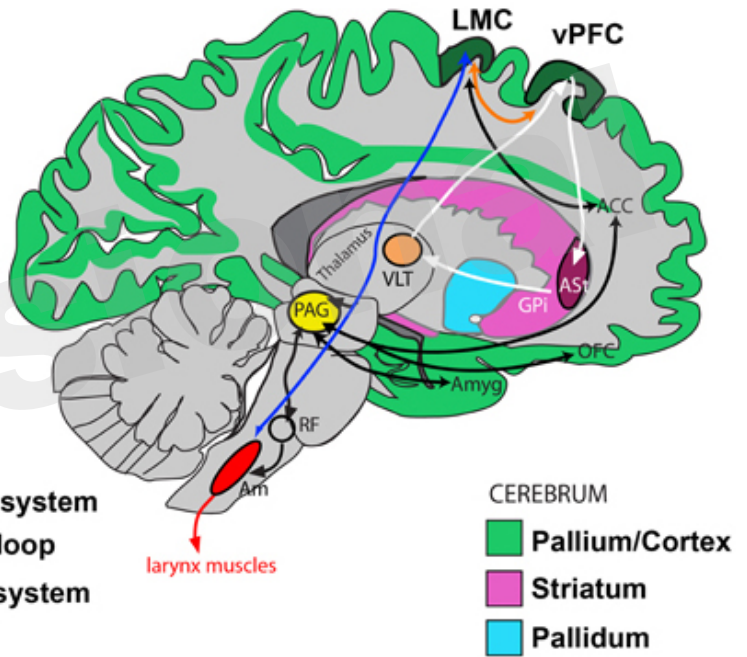


2

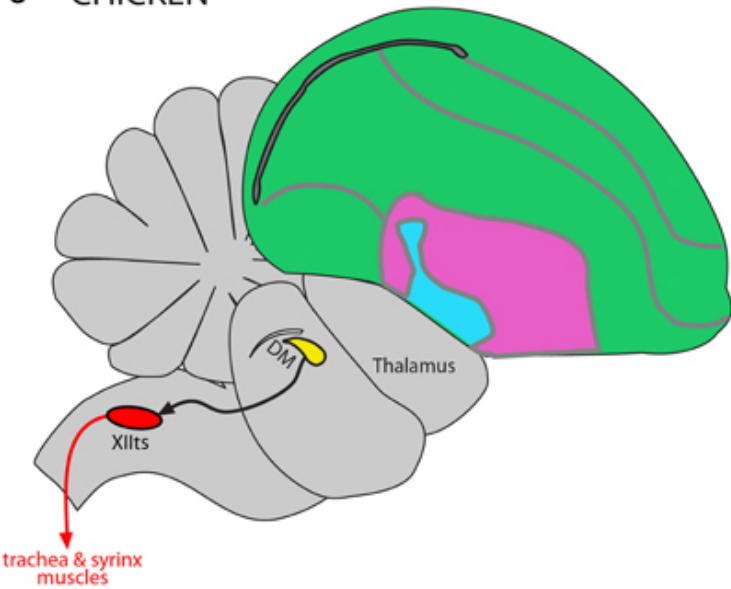
A SONGBIRD



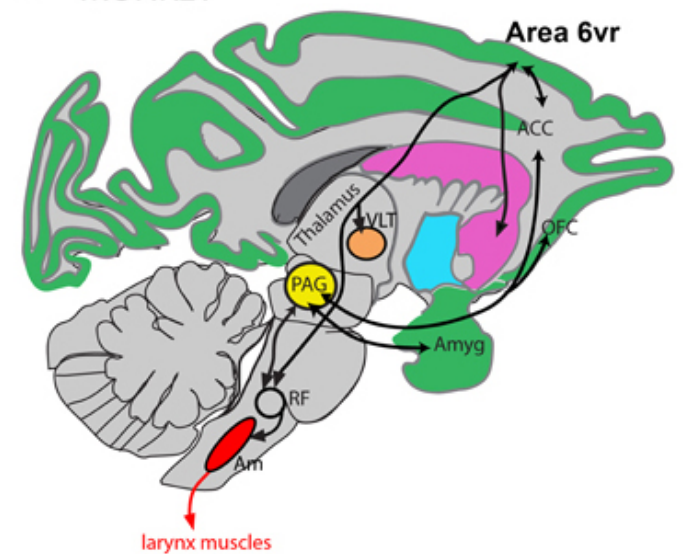
B HUMAN



C CHICKEN



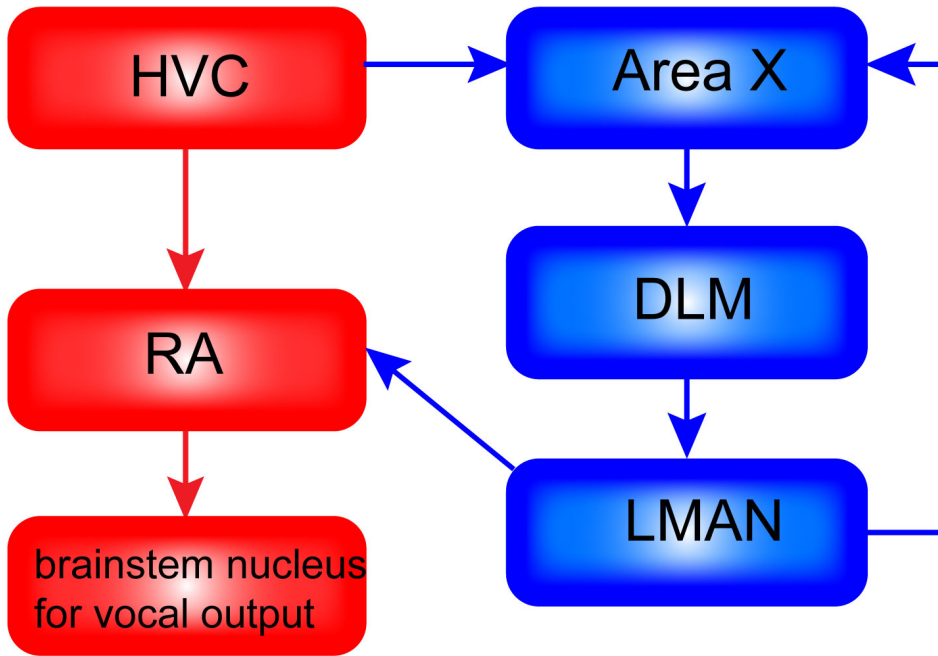
D MONKEY



A: Vocal pathways Figure 2.JPEG

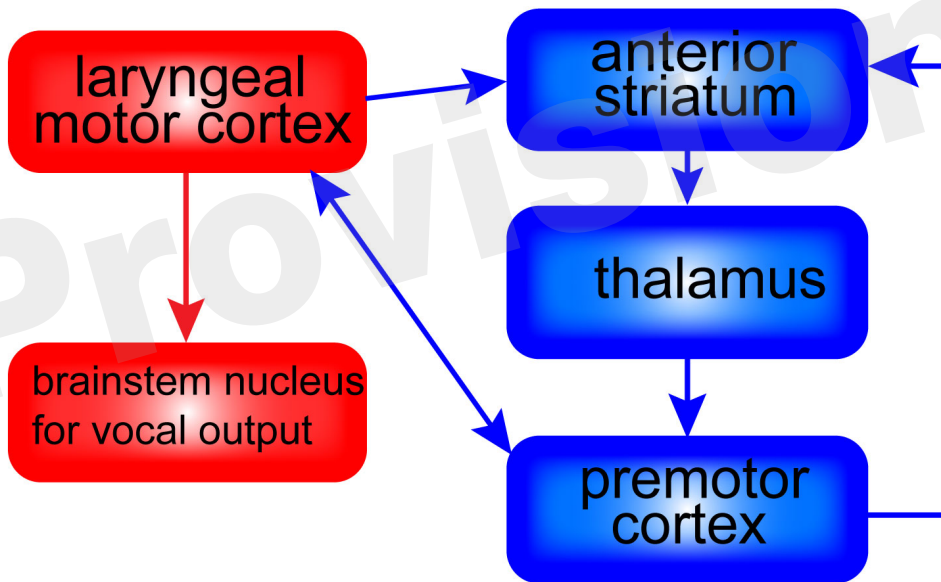
i: Songbirds

vocal motor pathway vocal learning pathway



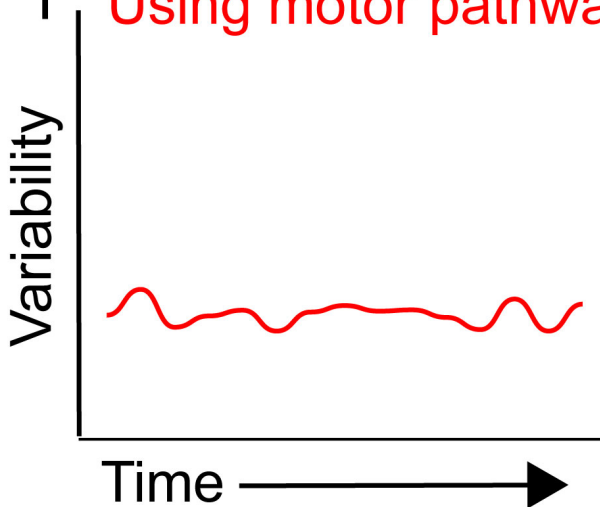
ii: Humans

vocal motor pathway vocal learning pathway



B: Vocal variability

i: Using motor pathway



ii: Using learning pathway

