BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

COMPUTATIONAL MODELING OF THE

NEURAL SUBSTRATES OF STUTTERING AND INDUCED FLUENCY

by

OREN CIVIER

B.A, Open University of Israel, 2000

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2010

Approved by

First Reader   _____
Frank H. Guenther, Ph.D.
Professor of Cognitive and Neural Systems
Boston University

Second Reader   _____
Daniel H. Bullock, Ph.D.
Professor of Cognitive and Neural Systems and Psychology
Boston University

Third Reader   _____
Ludo Max, Ph.D.
Associate Professor of Speech and Hearing Sciences
University of Washington

Dedicated to my late grandmothers, *safta* Rivka and *safta* Haya,

who waited for me to graduate before leaving this world

ACKNOWLEDGEMENTS

chapter of this dissertation; and Hayo Terband, who I know for less than a year, but whose support and friendship were crucial in the months leading to my defense.

Before concluding, I wish to thank Steve Grossberg for accepting me to the department, and guiding me in my first year of studies. Without his faith in me, I would probably not be where I am today. I am also grateful to my family for encouraging me to continue my studies, and for giving me emotional and financial support. I cannot forget the smile on the faces of my late grandmothers when they heard of my graduation, it is sad that they could not share with me the happiness for much longer. Lastly, I want to thank from the depths of my heart my soon-to-become-wife, Maria de los Angeles Vilte. Her help with the manuscript pales in comparison to her physical and mental support: she cared for me, fed me, entertained me, hugged me, and at times, also cried with me. I could not have made it without her unconditional love.

COMPUTATIONAL MODELING OF THE

NEURAL SUBSTRATES OF STUTTERING AND INDUCED FLUENCY

(Order No.   )

OREN CIVIER

Boston University Graduate School of Arts and Sciences, 2010

Major Professor: Frank H. Guenther, Professor of Cognitive and Neural Systems

ABSTRACT

Stuttering is a speech motor control disorder of unknown etiology whose hallmark is part-syllable repetition. The principal aim of this dissertation is to understand the neural mechanisms underlying stuttering through computational modeling with DIVA and GODIVA, neural network models of speech acquisition and production.

The first part of the dissertation investigates the hypothesis that stuttering may result in part from impaired readout of feedforward commands for speech, which forces persons who stutter (PWS) to produce speech with a motor strategy that is weighted too much toward auditory feedback control. Over-reliance on feedback control leads to sensory errors which, if they grow large enough, can cause the motor system to "reset" and repeat the current syllable. This hypothesis is investigated by impairing the feedforward control subsystem of the DIVA model. The model's outputs are compared to published acoustic data from PWS' fluent speech, and to combined acoustic and articulatory-movement data collected from the dysfluent speech of one PWS. The

simulations mimic the errors observed in the PWS subject's speech, as well as the repairs of these errors. Additional simulations were able to account for enhancements of fluency gained by slowed/prolonged speech and masking noise.

The second part of the dissertation explores the role of the basal ganglia (BG) – left ventral premotor cortex (vPMC) loop in the impaired readout of feedforward control. Two hypotheses are put to test: 1) due to structural abnormality in the corticostriatal projections carrying corollary discharge of motor commands, the BG fail to detect the context for initiating the next syllable, and 2) due to increased dopamine binding in the striatum leading to a ceiling effect, the BG are unable to bias cortical competition in favor of the correct syllable. Simulations of a neurally impaired version of the extended GODIVA model show that both hypotheses can explain dysfluent speech and associated abnormal brain activations. Further simulations account for the alleviation of stuttering with D2 antagonists. Together these results support the hypothesis that many dysfluencies in stuttering are due to abnormalities interfering with normal BG-vPMC loop operation, which ultimately biases the system away from feedforward control and toward feedback control.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BG……………………………………….. Basal Ganglia

BOLD……………………………………… Blood Oxygenation-Level-Dependent

CNS………………………………………… Central Nervous System

DA…………………………………………… Dopamine

DIVA……………………………………….. Direction Into Velocities of Articulators

F1…………………………………………... First formant

F2………………………………………….... Second formant

F3……………………………………….. Third formant

fMRI……………………………………….. functional Magnetic Resonance Imaging

GODIVA…………………………………... Gradient Order (or Gated Onset) DIVA

GPe…………………………………………… external Globus Pallidus

GPi…………………………………………. internal Globus Pallidus

IFS………………………………………….. Inferior Frontal Sulcus

LPC………………………………………… Linear Predictive Coding

PNS………………………………………… Persons who do Not Stutter

PWS……………………………………….... Persons Who Stutter

SMA………………………………………... Supplementary Motor Area

vMC………………………………………… ventral primary Motor Cortex

vPMC……………………………………… ventral PreMotor Cortex

WMF………………………………….......... White Matter Fibers

1 INTRODUCTION

Stuttering affects tens of millions of persons around the globe. It is characterized by frequent repetitions of syllables or parts of syllables (sound/syllable repetitions), by prolongations of sounds, and by blocks of speech. Persons who stutter (PWS) often describe the disorder as a loss of control and inability to proceed with their speech. Stuttering has probably received more attention than any other speech disorder (Van Riper, 1982), but its etiology remains elusive. Nevertheless, we now know much more about stuttering than 100 years ago. It is now clear that stuttering is a developmental disorder, that it has a genetic basis, and that it affects more males than females (Bloodstein & Ratner, 2008). Stuttering also relates to emotional state, and can be temporarily alleviated by various fluency enhancers such as choral speech (Andrews et al., 1983; Starkweather, 1987). Although this voluminous knowledge about stuttering cannot yet be translated into effective means of treatment (Curlee, 1999), it does strongly suggest that stuttering is a neurologically-based disorder (Max, 2004; Max, Guenther, Gracco, Ghosh, & Wallace, 2004). This is the starting point of this dissertation.

1.1 Hypothesis

Recent findings of neural abnormalities in the brains of PWS (Chang, Erickson, Ambrose, Hasegawa-Johnson, & Ludlow, 2008; Sommer, Koch, Paulus, Weiller, & Buchel, 2002; Watkins, Smith, Davis, & Howell, 2008; Wu et al., 1997) provide the best evidence for a neurological basis of stuttering. The main hypothesis investigated in this dissertation is that some or all of these neural abnormalities lead to an impairment in the ability of PWS to read out motor commands for well-learned syllables (*feedforward*

commands). This impairment has several possible consequences; each leads to a different type of dysfluency. If no motor commands are read out at all, the outcome is a block. If the commands to start voicing are read out, but the commands to move the articulators are not, the outcome is a prolongation. If commands are improperly read out (either partial or weak readout), speech production continues, but production errors prevail. The hypothesis further assumes that to correct these errors PWS use feedback-based speech motor control or *feedback commands*. These commands can correct some errors, but sooner or later the error on the current syllable accumulates, and PWS are forced to halt speech production and repeat the erroneous syllable, thus producing a sound/syllable repetition.

The goal of this dissertation is to test the aforementioned hypothesis, to generate predictions based on this hypothesis, and to use the hypothesis to account for several fluency enhancers: slowed/prolonged speech, masking noise, and treatment with D2 antagonists. Insights into the neurological basis of stuttering may help refine current treatment methods, and develop new ones. Finally, our understanding of normal speech production should improve as well. Accounting for stuttering due to impairment in a given region of the brain can clarify that region's function in fluent speech.

1.2   Neurocomputational modeling

Until the 1990s, not much was known about brain function in stuttering, or even in speech in general. Speech is a behavior restricted to humans, and as such, neurophysiological recordings are only available in rare circumstances. The recent advent of functional brain imaging techniques has led, however, to an accumulation of data

regarding abnormal brain activations associated with stuttering (for review, see Bloodstein & Ratner, 2008; De Nil, 2004; Ingham, 2001, 2008). As these data accumulate, it is becoming increasingly important to have a modeling framework within which to interpret them. To be testable, such a model of stuttering should be also capable of making predictions that can be experimentally confirmed or disconfirmed. Unfortunately, none of the existing models of stuttering fit these criteria. Many models of stuttering do not describe brain functions (Harrington, 1988; Kalveram, 1991, 1993; Neilson & Neilson, 1987; Zimmermann, 1980b), and the more recent models that do relate to the brain (Giraud et al., 2007; Ingham, Ingham, Finn, & Fox, 2003; Wu et al., 1995) are box-and-arrow models, which cannot make quantitative predictions.

The DIVA (Directions Into Velocities of Articulators, see Guenther, Ghosh, & Tourville, 2006) and GODIVA (Gradient Order DIVA, see Bohland, Bullock, & Guenther, in press) models of speech acquisition and production overcome the limitations of prior models. Being neurobiologically plausible, both models can be used to interpret functional imaging data, and being computational, both can make quantitative predictions (Guenther, 2006; Guenther et al., 2006; Tourville, Reilly, & Guenther, 2008). Furthermore, because these models drive an artificial articulator (Maeda, 1990), they can also fit articulatory movement and acoustic data (Guenther, 1995, 2006; Guenther et al., 2006; Guenther, Hampson, & Johnson, 1998; Nieto-Castanon, Guenther, Perkell, & Curtin, 2005; Tourville et al., 2008). In this dissertation, the DIVA and GODIVA models are used to account for abnormal brain activations associated with stuttering, as well as for articulatory movements and acoustics during the fluent and dysfluent speech of PWS.

The DIVA model is also used to fit data on the frequency of speech dysfluencies that PWS generate in different speech conditions.

The DIVA and GODIVA models address different aspects of speech production. The DIVA model deals with motor control for speech articulation ("low-level" speech production), while the GODIVA model deals with speech planning and initiation (at the intersection between "low-level" and "high-level" speech production, see Bohland, 2007). To have a modeling framework capable of addressing stuttering from the hypothesized (internal) neural abnormality up to the (external) behavioral manifestation, the two models must be integrated. In the current implementation the GODIVA model communicates with the DIVA model through the left ventral premotor cortex (vPMC) that consists of speech sound map (SSM) cells. The GODIVA model decides what SSM cell should be active at each point, and the DIVA model executes the motor program read out by that cell. While executes the commands of the program, the DIVA model is continuously updating the GODIVA model of its progress[1].

To account for stuttering, the DIVA and GODIVA models need to be extended. A monitoring subsystem is added to the DIVA model (Section 2.2.2), and a basal ganglia (BG) - vPMC loop module is added to the GODIVA model (Section 3.2.2). Although used here to account for stuttering, both neural circuits have important roles in normal speech production: the monitoring subsystem detects and repairs speech errors that often

---

[1] The DIVA model updates the GODIVA model of its progress by sending a corollary discharge (or a copy) of the motor command currently being executed. The GODIVA model uses this information to predict when the syllable is about to terminate.

occur in fluent speech, and the BG-vPMC loop facilitates production of rapid sequences of syllables.

1.3   Organization of the dissertation

Chapters 2 and 3 contain the body of the research. Chapter 2 tests the part of the hypothesis that accounts for generation of repetitions due to impairment in the readout of feedforward control, and the resulting bias toward feedback control. Because feedforward and feedback control are "low-level" aspects of speech production, simulations with the DIVA model are sufficient. Furthermore, since simulations of the DIVA model are easily compared to articulatory movement and acoustic data, Chapter 2 includes no comparison to imaging data. Chapter 3 tests the hypothesis as a whole and therefore uses both the GODIVA and DIVA models. In that chapter the simulations are compared primarily to functional imaging data. Chapter 4 concludes the dissertation. It discusses the insights into the treatment of stuttering derived from this work, and suggests future directions for related research.

2   OVERRELIANCE ON AUDITORY FEEDBACK MAY LEAD TO SOUND/SYLLABLE REPETITIONS

2.1   Introduction

In his classic book "The Nature of Stuttering," Van Riper (1982) describes a person that stopped stuttering "after an incident in which he became completely deafened. The cessation of stuttering occurred within three hours of the trauma and shortly after he began to speak" (p. 383–384). What is the relationship between sensory feedback (in this case, hearing oneself) and stuttering? Although numerous investigations have attempted to account for the effect of various fluency enhancers involving altered sensory feedback (e.g., Fairbanks, 1954; Hutchinson & Ringel, 1975; Loucks & De Nil, 2006a; Mysak, 1960; Neilson & Neilson, 1987, 1991; Webster & Lubker, 1968) no account has received overwhelming support. The obstacle may lie in the classic experimental devices typically used to study stuttering rather than the theoretical models themselves.

Most researchers have used psychophysical experiments to investigate sensory feedback in speech, but the results are often difficult to interpret because several feedback channels are simultaneously active during a typical experiment. In addition, experimental blockage of proprioceptive feedback channels is unfeasible (Scott & Ringel, 1971, p. 806), and auditory blockage is incomplete (e.g., Adams & Moore, 1972). Alternatively, sensory feedback can be studied by using computational models of speech production where feedback channels can be systematically blocked or altered.

In this chapter we use the DIVA model (Directions Into Velocities of Articulators, see Guenther et al., 2006), which is a biologically plausible model of speech production

that mimics the computational and time constraints of the central nervous system (CNS), to test our hypothesis regarding the involvement of sensory feedback in stuttering. The DIVA model differs from other computational models applied to stuttering (e.g., Kalveram, 1991, 1993; Neilson & Neilson, 1987; Toyomura & Omori, 2004) in its ability to simulate the kinematics and acoustics of both dysfluent and fluent speech (for DIVA simulations of fluent speech see Guenther, 1995; Guenther, 2006; Guenther et al., 2006; Guenther et al., 1998; Nieto-Castanon et al., 2005), which makes it possible to compare the simulations to a much larger set of data than permitted by other models.

Numerous authors have suggested that stuttering may be due in part or whole to an aberrant sensory feedback system. Some have hypothesized that persons who stutter (PWS) differ from those who do not stutter (PNS) by relying too heavily on sensory feedback (De Nil, Kroll, & Houle, 2001, p. 79; Hutchinson & Ringel, 1975; Tourville et al., 2008, p. 1441; van Lieshout, Peters, Starkweather, & Hulstijn, 1993), while others have claimed that PWS actually benefit from reliance on sensory feedback (Max, 2004; Max et al., 2004; van Lieshout, Hulstijn, & Peters, 1996a, 1996b). Our hypothesis is that due to an impaired feedforward control system, PWS rely more heavily on a feedback-based motor control strategy which has the consequence of increasing the frequency of motor errors (cf. Jäncke, 1991; Kalveram, 1991; Kalveram & Jäncke, 1989; Lane & Tranel, 1971; Stromsta, 1972; 1986, p. 204; Toyomura & Omori, 2004; Van Riper, 1982, p. 387; Zimmermann, 1980b). Compared to stored feedforward commands, which can be read out directly from memory, feedback control requires the detection and correction of production errors and involves significant lags due to the time it takes to detect and

process these errors (e.g., Guenther et al., 2006). As the amount of error accumulates, it is

our contention that this eventually leads to a motor "reset" in which the system attempts

to restart the current syllable. Such a reset would constitute a sound/syllable repetition[2].

Forty years ago, Wingate (1969**,** p. 107; 2002**,** p. 327) suggested that moments of

stuttering result from errors in the articulatory movements required to transition between

phonemes; a hypothesis that found support in follow-up experimental work (Adams,

1978; Agnello, 1975; Harrington, 1987; Howell & Vause, 1986; Stromsta, 1986**,** p. 94;

Stromsta & Fibiger, 1981; Zimmermann, 1980b). In an attempt to clarify this mechanism,

Van Riper (1982**,** p. 117, 435) and Stromsta (1986**,** p. 14, 28, 91, 111) reasoned that

following an error, the CNS detects a problem and searches for the response capable of

repairing it; the response is often a "reset" in which the current syllable is attempted

again. Postma and Kolk (1993**,** p. 482) went further, combining the speech motor control

approach to stuttering with Levelt's (1983) self-repair theory. By assuming that speech

motor execution can be monitored via sensory feedback, they argued that "motor errors

are supposed to be detectable, and when reacted to, would cause stuttering". To

emphasize that motor errors can be detected via sensory feedback (Guenther, 2006;

Postma, 2000) we use the term *sensorimotor errors* (or simply *errors*).

Here we investigate the hypothesis that, due to impairment of the ability to

properly read out feedforward commands, PWS are forced to rely too heavily on auditory

---

[2] We use the term sound/syllable repetition (Conture, 2001**,** p. 6; Wingate, 1964) to refer to any audible repetition of a syllable or part of it, without regard to phonemic boundaries (the cut-off can be within or between phonemes).

feedback control[3], which can lead to an accumulation of error on the current production. The repairs of these errors take the form of sound/syllable repetitions (or *repetitions*). The idea was presented in a limited form previously by our group (Civier & Guenther, 2005; Max et al., 2004) and inspired a series of simulations aimed at showing that when the DIVA model is biased away from feedforward control and toward feedback control, it behaves similarly to PWS, both in the type of errors produced, and in the way the errors are repaired. This demonstrates that over-reliance on auditory feedback (due to impairment of feedforward commands) may indeed be the reason for the repetitions made by PWS.

This chapter is organized as follows. After a description of the DIVA model, we report a series of modeling experiments in which the model is biased away from feedforward control and toward feedback control. The first modeling experiment mimics sensorimotor errors (observable in the acoustic domain) produced by PWS during fluent speech production. The second experiment tests whether the DIVA model's attempt to repair these errors resemble the acoustic and kinematic properties of sound/syllable repetitions produced by a PWS. The third and fourth experiments examine the model's ability to account for the reduction in the frequency of repetitions in two fluency-inducing conditions: slowed/prolonged speech and masking noise. To conclude we suggest how to further test our hypothesis using the model.

---

[3] The impairment in the readout of feedforward commands from motor memory may be due, for example, to basal ganglia problems (cf. Alm, 2004). The current investigation does not speak directly to the cause of the feedforward impairment; instead the focus is on the consequences of biasing away from feedforward control and toward feedback control. We also acknowledge that in addition to auditory feedback control, the biasing may be toward somatosensory feedback control as well (as suggested by De Nil et al., 2001; Hutchinson & Ringel, 1975; van Lieshout et al., 1996a, 1996b; van Lieshout et al., 1993). The discussion and simulations in this chapter, however, are limited to the auditory feedback channel.

2.2   Neurocomputational modeling of speech kinematics and acoustics

The DIVA model is a biologically plausible neural network model capable of simulating the production of fluent speech and its development, e.g., learning of speech sounds in the babbling phase (e.g., Callan, Kent, Guenther, & Vorperian, 2000; Guenther, 1994; Guenther, 1995; Guenther et al., 1999; Guenther et al., 2006; Guenther et al., 1998; Nieto-Castanon et al., 2005; Perkell, Guenther et al., 2004; Perkell, Matthies et al., 2004). The model combines mathematical descriptions of underlying commands with hypothesized neural substrates corresponding to the model's components. The model is implemented in computer simulations in which it controls the movements of articulators in an articulatory synthesizer (Maeda, 1990). In the model, which is schematically represented in Fig. 2-1, cells in the motor cortex generate the overall motor command, M(t), to produce a speech sound. M(t) is a combination of a *feedforward command* and a *feedback command*:

$$\dot{M}(t) = \alpha_{ff}\dot{M}_{feedforward}(t) + \alpha_{fb}\dot{M}_{feedback}(t)$$

where α_{ff} and α_{fb} represent the amount of feedforward and feedback control weighting, respectively. The default parameter values are $\alpha_{ff} = 0.85$ and $\alpha_{fb} = 0.15$, settings that can account for normal speech, as well as for speech produced during auditory feedback perturbation (Tourville et al., 2008).
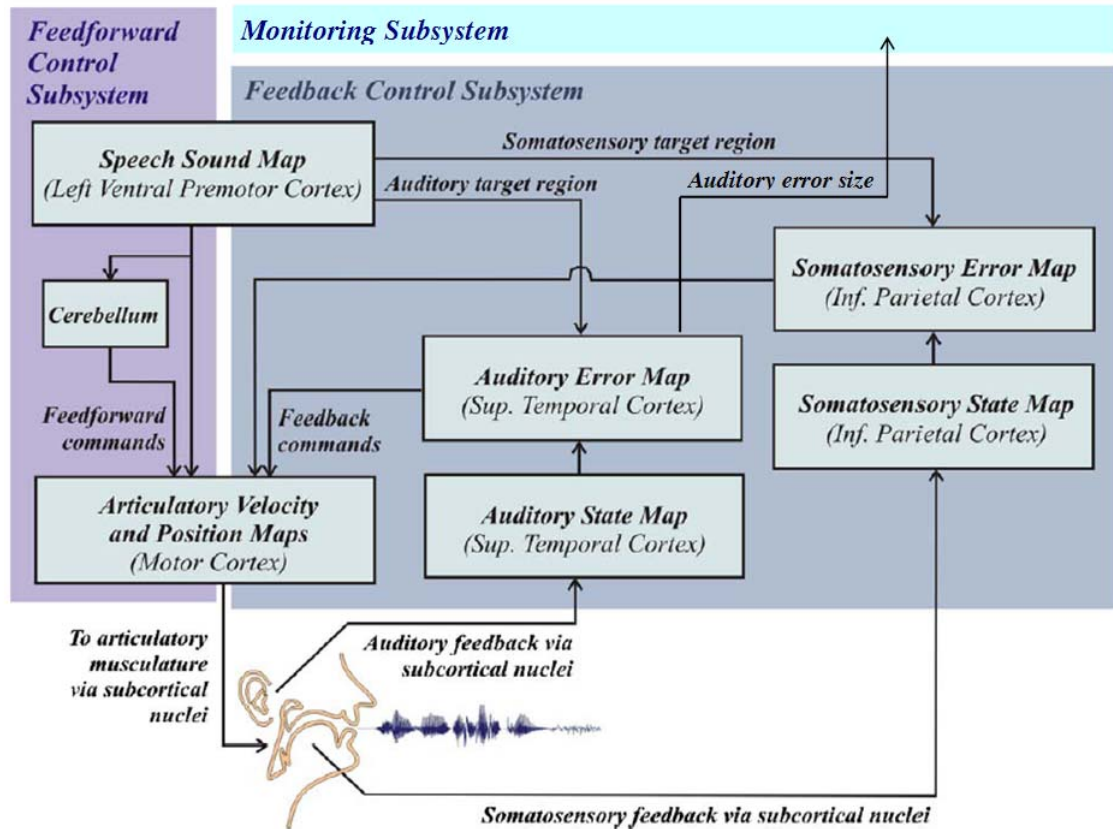
Fig. 2-1. Schematic of the DIVA model. Each box corresponds to a set of neurons (or map) in the model, and arrows correspond to synaptic projections that transform one type of neural representation into another. The model is divided into three distinct systems: the Feedforward control subsystem on the left, the Feedback control subsystem on the lower-right, and the Monitoring subsystem on the upper-right (reproduced from Guenther et al., 2006, with the exception of the Monitoring subsystem).

2.2.1 Feedback and Feedforward control subsystems

Feedback, or closed-loop control (Borden, 1979; Max et al., 2004; Smith, 1992), relies on sensory feedback, and therefore, can adapt to error-inducing perturbations to the speech system, such as the introduction of a bite block (e.g., Baum, McFarland, & Diab, 1996; Hoole, 1987), unexpected jaw perturbation (Guenther, 2006), and altering of auditory feedback of one's own speech (e.g., Tourville et al., 2008). In the *feedback control*

subsystem of the DIVA model, auditory and somatosensory feedback loops (Guenther, 2006) are constantly comparing speech output with the desired auditory and somatosensory targets, which consist of time-varying regions that encode the allowable variability of the acoustic and somatosensory signal throughout the utterance (the upper and lower bounds of a target region are tuned when an infant listens to examples of a speech sound, as described in Guenther et al., 2006). When a mismatch (or *error*) is detected in one of the loops, corrective motor commands which were acquired during the babbling phase of the model, are generated in order to eliminate the error. In a sense, speaking using only auditory feedback control is like trying to sing along to a tune playing on the radio for the first time – there will always be a delay between what the listener hears and what he produces. At each point of time, therefore, there will be a mismatch between the played tune (analogous to the auditory target in the DIVA model) and the tune the listener hears himself producing (analogous to the auditory feedback in the DIVA model).

To prevent errors that inevitably occur when using feedback control alone, speakers also utilize open-loop or feedforward control mechanisms (Borden, 1979; Kent, 1976; Kent & Moll, 1975; Lashley, 1951; Max et al., 2004; Smith, 1992). The *feedforward control subsystem* of DIVA issues preprogrammed commands that were acquired through prior attempts at producing the target utterance. These commands do not rely on error detection via sensory feedback and therefore avoid the time-lap problems of feedback control. Speaking with feedforward control is analogous to singing a song from memory. The strongest evidence for feedforward control during speech

production comes from a variety of reports that illustrate that speech motor behavior is quite resistant to temporary loss in somatosensory information (Abbs, 1973; Gammon, Smith, Daniloff, & Kim, 1971) as well as temporary (Gammon et al., 1971) and permanent (Cowie & Douglas-Cowie, 1983; Goehl & Kaufman, 1984) loss of auditory information. Finally, feedforward control is also responsible for the after-effects observed after the sudden removal of speech perturbations that were applied for extended periods of time (e.g., Purcell & Munhall, 2006; Villacorta, Perkell, & Guenther, 2007)

2.2.2 Error maps and the Monitoring subsystem

During normal speech acquisition and production the *auditory* and *somatosensory error maps* (each map, or set of neurons, is represented by a box in Fig. 2-1) have multiple roles. Initially, they assist in learning feedforward commands so the system does not have to constantly rely on feedback to reliably produce speech. Later, during running speech, they calculate error between desired and actual sensory consequences, which can arise when the speech system is unexpectedly perturbed. Finally, we hypothesize that auditory errors are inspected by a *monitoring subsystem* (not previously described by the DIVA model; see Fig. 2-1) that is responsible for detecting and repairing errors that are too large to be corrected by feedback control loops. Errors are detected by inspecting auditory feedback but not somatosensory feedback following Levelt's (Levelt, 1983, 1989) implementation of a speech monitor (perceptual loop theory); a theoretical design consideration that has some empirical support (Postma, 2000). Levelt (1989) limited the scope of his model to language formulation errors, but it may be extended to deal with non-phonetic errors as well (Postma, 2000).

The monitoring subsystem uses the following equation in order to calculate the probability of an error repair, R_{error}, being triggered at time t:

$$P\left(R_{error}(t)\right) = \varepsilon * 10^{-3} * E_{auditory}(t)$$

where the constant ε is set to 2, t is time in milliseconds, and $E_{auditory}(t)$ is the auditory error, measured in Hz, calculated by auditory error cells in the model's superior temporal cortex.

The equation is probabilistic to account for the fact that not all speech errors are repaired (e.g., Levelt, 1983, p. 59). The probabilistic occurrence of repairs may explain the variability of stuttering due to emotions, muscle tension, and other perceptual and physiological factors (Alm, 2004, p. 332, 344, 359; Van Riper, 1982; Zimmermann, 1980b, p. 122, 133). While a probabilistic function cannot capture the nuances of the factors affecting stuttering, it may at least imitate the unpredictable nature of the disorder. According to the equation, repair occurrence depends on error size as well (cf. Zimmermann, 1980b, p. 131). This design consideration is supported by data showing that stutter occurrence can be predicted by the extent of sensorimotor error, measured as lack of anticipatory coarticulation (Stromsta, 1986, p. 79; Stromsta & Fibiger, 1981).

Once a repair is triggered, the monitoring subsystem will be also responsible to execute it. The implementation of repairs in the DIVA model is described in Section 2.4.1.

2.3   Modeling experiment 1: Bias toward feedback control leads to sensorimotor

errors

There is evidence that PWS have errors not only in their dysfluent speech preceding repetitions, but in their fluent speech as well (Bloodstein, 1995**,** p. 35; Postma & Kolk, 1993**,** p. 482; Robb & Blomgren, 1997; Stromsta, 1986**,** p. 189, 191; Van Riper, 1982**,** p. 396, 402). As noted earlier, a possible cause of these errors is PWS' impaired ability to read out feedforward commands resulting in an overreliance on auditory feedback, which can lead to delays due to time lags associated with afferent information (Guenther et al., 2006; Neilson & Neilson, 1987, 1991). The DIVA model can be biased away from feedforward control and toward feedback control by utilizing a relatively low gain on the feedforward control subsystem in conjunction with relatively high gain on the feedback control subsystem.

Without sufficient feedforward control, sensorimotor errors arise, and the auditory feedback loop detects them as auditory errors. Consistent with CNS physiological delays, the DIVA model takes 20ms to perceive auditory feedback, and then an additional 42ms to execute the corrective motor commands (Guenther et al., 2006); but since some consonant sequences contain articulatory events separated by as little as 10ms (Kent, 1976; Kent & Moll, 1975), the DIVA model is expected to be too slow in issuing the corrective feedback commands and thus unable to correct the errors in a timely manner. The errors will then grow large, increasing the likelihood of a motor "reset". Nevertheless, in this modeling experiment we only report of simulation instances where

none of the errors triggered a stutter. Fluent instances can be compared to data on PWS'

fluent speech, which are more abundant than data on their dysfluent speech.

Errors due to reliance on auditory feedback are expected to be the greatest in

formant transitions, especially if the transitions are rapid. This is anticipated from a

system biased away from feedforward control and toward feedback control, since during

rapid transitions, such as the ones required for consonants (Borden, 1979), there are

substantial mismatches between the fast changing target region formants, and the slow

changing formants the feedback controller is capable of producing. The prominence of

errors on rapid transitions can also be observed in the results of auditory tracking

experiments where auditory feedback is heavily used — the largest errors follow rapid

shifts in the auditory target (e.g., Nudelman, Herbrich, Hoyt, & Rosenfield, 1987).

2.3.1   Methods

To bias the DIVA model away from feedforward control and toward feedback control,

the feedforward and feedback gain parameter values were modified to $\alpha_{ff} = 0.25$ and

$\alpha_{fb} = 0.75$ (instead of the default settings where $\alpha_{ff} > \alpha_{fb}$). Simulations were then

performed with both the modified settings (*stuttering DIVA version*) and the default

settings (*non-stuttering DIVA version*). To assess performance on different formant

transition speeds, we used the utterance /bid/ which requires rapid second (F2) and third

(F3) formant transitions, and the utterance /bad/ that requires slower transitions (this can

be seen by comparing the slopes of the formant transitions from /b/ to /i/ in Fig. 2-2(a)

with the slopes of the transitions from /b/ to /a/ in Fig. 2-2(c)). Each simulation was repeated until the DIVA model produced the utterance fluently.

In all experiments, auditory errors were based on formant values. For each one of the first three formants, a *formant error* was determined, which is defined as the mismatch (in Hz) between the actual formant values for that production and the target formant values at each time interval within the syllable/word. In addition to individual F1, F2 and F3 formant errors, we also report the sum of the formant errors: the (total) *auditory error* (or $E_{auditory}$, as defined in Section 2.2.2). If, within a given production, all three formants are within their corresponding target regions, auditory error at that time point is 0 Hz (e.g., Fig. 2-2(a) at 158ms).

2.3.2   Results

Fig. 2-2 shows the target region and auditory feedback of the non-stuttering (left plots) and stuttering (right plots) versions of the DIVA model during a fluent production of both /bid/ (upper plots) and /bad/ (lower plots). To illustrate the performance of the model during the voiced (white background) as well as the silent (shaded background) periods of the utterance, these data are plotted as if voicing is always on. To emphasize, however, that auditory feedback is not available when voicing is off, the auditory errors are plotted in the voiced periods only.

(a) /bid/ - non-stuttering version

(b) /bid/ - stuttering version

(c) /bad/ - non-stuttering version

(d) /bad/ - stuttering version
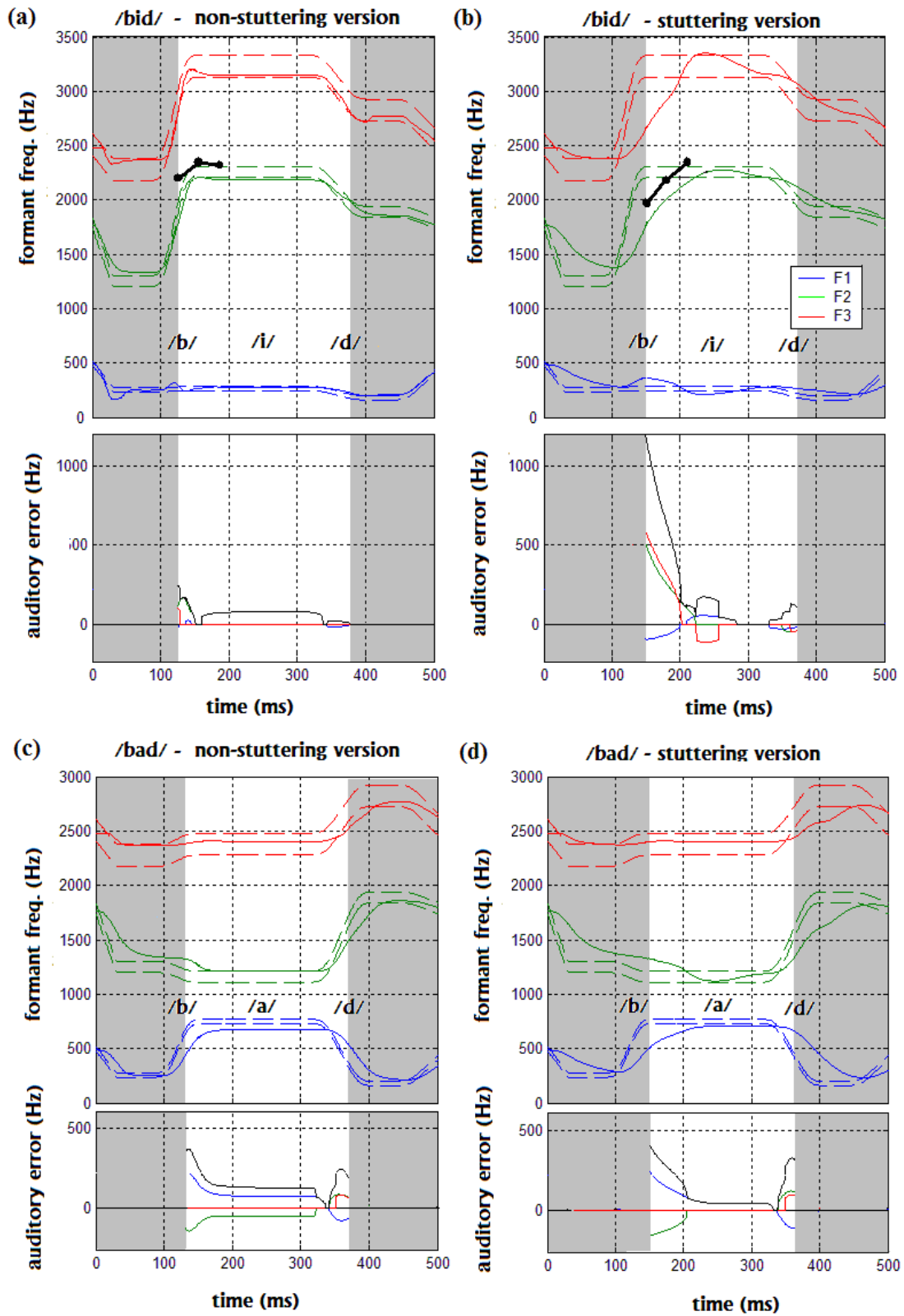
Fig. 2-2. Auditory feedback, target region, and errors for the non-stuttering (left plots) and stuttering (right plots) versions of the DIVA model, producing /bid/ (top plots) and /bad/ (bottom plots). The top of each plot contains the target region lower and upper bound frequencies (dashed lines), with one frequency pair for each formant (different formants coded by different colors). For each formant, the auditory feedback frequency during the simulation is indicated by solid colored line (including in the shaded periods were voicing was off, see Section 2.3.2 for details). The extent of the total auditory error (black) and the formant auditory errors (other colors) are plotted on the bottom of each plot. Indicated on the top plots are the results from Robb and Blomgren (1997) who sampled F2 during productions of /bid/ at 0ms, 30ms, and 60ms after voicing onset (black dots connected by a line). Voicing onset follows initial articulatory movement by 28.88ms for PWS and 18.10ms for PNS (Healey & Gutkin, 1984).

Note that the stuttering version, which relies more on feedback control, had larger errors than the non-stuttering DIVA version, and that the largest error occurred during formant transitions. Moreover, the errors made by the stuttering DIVA version on /bid/ (rapid transition) exceeded 1000Hz — much larger than the errors made by the stuttering DIVA version on /bad/ (slow transition), which never rose above 500Hz.

2.3.3 Discussion

The simulation results demonstrate that a bias toward feedback control leads to increased auditory error, especially on utterances with rapid formant transitions.

Do PWS show similar acoustic patterns during fluent speech production? There are repeated reports that PWS exhibit abnormal F2 transitions during fluent speech (Robb & Blomgren, 1997; Subramanian, Yairi, & Amir, 2003; Yairi & Ambrose, 2005, p. 259). The data most relevant to our simulation comes from a pair of studies (Blomgren, Robb, & Chen, 1998; Robb & Blomgren, 1997) who studied F2 transitions and steady states in

CVC words produced by PWS and PNS[4]. Blomgren et al.'s finding that PWS had greater vowel centralization than PNS suggests, much like our results do, that PWS have larger auditory errors. Moreover, PWS produced /Cit/ and /Cat/ syllables differently; compared to PNS, PWS had significantly lower F2 in the steady state portion of /Cit/ syllables only. To compare these results with our simulations, we approximated the F2 formant error (during the steady state portion of the syllable) for Blomgren et al.'s data by assuming that PNS perfectly reproduce the desired auditory target (have no errors). The F2 formant error of PWS is then equal to the difference in F2 steady state values between the two groups. For the study in question, the group differences in F2 steady state values were 125Hz for /bit/, and 94Hz for /bat/. These results are qualitatively similar to the stuttering DIVA version's steady state F2 formant errors[5]: 127Hz on /bid/, and 68Hz on /bad/[6].

Other studies that examined steady state formant values are consistent with our simulations as well. For example, data from several studies suggest, as Blomgren et al.'s (1998) data do, that PWS have significant F2 errors on /CiC/ syllables (Hirsch et al., 2007; Klich & May, 1982; Prosek, Montgomery, Walden, & Hawkins, 1987). One of the studies (Prosek et al., 1987) also affirms Blomgren et al.'s finding that there is no

---

[4] Comparison of the simulations with the kinematics of PWS' fluent speech is not included because most studies that measured the tongue movements of PWS (the tongue is the articulator whose position correlates with F2 the most, see Stevens, 1998) either did not report data (e.g., Alfonso, 1991; McClean, Tasko, & Runyan, 2004) or only reported qualitative data (McClean & Runyan, 2000) from speech segments with rapid F2 transitions.

[5] Because F2 did not always come to a complete steady state during the simulation, the F2 values used to calculate the error were the average of F2 values measured in five equally spaced points spanning the simulated vowel production, starting 40ms after voicing onset. Blomgren et al. (1998, p. 1044-1045) used the same procedure in order to calculate steady state F2 values.

[6] That the syllables from Blomgren et al. (1998) and Robb and Blomgren (1997) study terminate with /t/, while the syllables uttered by DIVA terminate by /d/ should not make a difference on the transition from the initial /b/ to the middle vowel. We did not use /t/ for our simulations because DIVA cannot produce that sound.

significant difference between the F2 of PWS and PNS when producing /CaC/, thus, PWS have no F2 formant error in this case. In conclusion, the simulations can account for several findings that show that PWS make larger errors on /CiC/ syllables than on /CaC/ syllables. According to the current account, /CiC/ syllables are more likely to have errors because /i/ has the highest F2 of all vowels (~2250Hz), requiring rapid F2 transitions. By contrast, the vowel /a/ has a much lower F2 (~1150Hz), and thus requires less rapid F2 transitions from the surrounding consonants.

In both the simulation and the aforementioned studies, the average F2 value produced by PWS (or the stuttering DIVA version) for the vowel in /CiC/ syllables is lower than the F2 value of the target region. Since the transition from the initial consonant to /i/ is an upward transition, the error may be attributed to an incomplete transition. This would explain why both children and adults who stutter display smaller than normal formant transitions (Caruso, Chodzko-Zajko, Bidinger, & Sommers, 1994; Subramanian et al., 2003). Furthermore, it is possible that the trend is particularly evident in bilabial consonants (examine transition onset and offset data in Chang, Ohde, & Conture, 2002) because they tend to require large F2 transitions; large transitions are more difficult to complete when heavily relying on feedback control.

Close inspection of Fig. 2-2(b) reveals that the error made by the stuttering version of DIVA on /bid/ is in large part due to a delayed F2 transition. This is a direct consequence of the bias toward feedback control; a control mechanism with inherent time lags. This pattern is similar to that reported by Robb and Blomgren (1997), which using the same subjects as Blomgren et al. (1998), provided evidence that PWS exhibited

delayed F2 transitions into the vowel as compared to the PNS group. For a comparison between our simulations and the Robb and Blomgren results, we have plotted the authors' data in Fig. 2-2. Specifically the median F2 at voicing onset, as well as 30ms and 60ms later, are plotted in Fig. 2-2(a), for the PNS, and in Fig. 2-2 (b) for the PWS. Note that the errors made by PWS appear to be associated with delays, as was the case for the errors made by the stuttering DIVA version.

Fig. 2-2(b) indicates that the errors made by the stuttering DIVA version on /bid/ are not only because the formant transitions started late, but also because the transitions themselves were slow (i.e., the transition slopes are flatter than those of the non-stuttering DIVA version depicted in Fig. 2-2(a)). This slowness can be attributed to the fact that the feedback control subsystem is limited in the speed of movement it can handle effectively (van Lieshout et al., 1996a, p. 89); faster transition speeds may be achieved by increasing the gain on the feedback control subsystem even more, but this can lead to undesirable overshoots and oscillations: the corrective motor commands will over-correct errors, generating new errors of the opposite sign (Ghosh, 2005; Stromsta, 1986, p. 203). Evidence for PWS' slowness is found in a report of F2 transition slopes for the syllables /beC/ and /meC/ whose transitions are similar to /bit/ in nature (Chang et al., 2002). In the age groups examined, children who stutter had slower transition rates in comparison with children who do not stutter[7].

---

[7] Robb and Blomgren (1997) reported F2 transition slopes as well. They reported F2 transition slope coefficients (the slopes of straight lines from F2 at voicing onset, to F2 at 30ms or 60ms later) for /bid/ and /bad/ produced by adult PWS. However, for the measurement to be accurate, transitions should be at least 30ms long; an assumption that may not always hold for the bilabial stop /b/ (Lieberman & Blumstein, 1988, p. 224; Miller & Baer, 1983).

2.4   Modeling experiment 2: Sensorimotor errors lead to sound/syllable repetitions

The correspondence between published data and the DIVA simulations in Experiment

One provides support for the view that the abnormal speech motor patterns observed in

the fluent speech of PWS could be the result of auditory errors that inevitably result from

a motor control strategy that relies too heavily on feedback. We also contend that

auditory errors, if large enough, can also cause sound/syllable repetitions commonly

observed in PWS. Recall from Fig. 2-1 that the DIVA model has a monitoring subsystem

that serves to repair auditory errors that are too large to be repaired by feedback control

only. The form in which the monitoring subsystem repairs the error is to restart the

current syllable until the auditory error is small enough to allow the forward flow of

speech. Therefore, repetitions result from the attempts to repair large sensorimotor errors.

In contrast to stuttering models that suggest a defect in speech monitoring — due to

hyper-sensitivity to errors (Martin, 1970; Russell, Corley, & Lickley, 2005; Sherrard,

1975), or due to inaccuracy in predicting the sensory consequences of movements (Max,

2004; Max et al., 2004) — within the DIVA model the problem of stuttering lies not in

erroneous monitoring subsystem, but in a faulty feedforward control system that forces

PWS to employ an error-prone feedback motor control strategy.

This general view that repetitions arise as a *self-repair* strategy is not new

(Postma & Kolk, 1993). However, previous attempts to apply the self-repair hypothesis

to stuttering have largely assumed linguistic (phonologic) formulation errors (Postma &

Kolk, 1993; Van Riper, 1971, 1982) rather than sensorimotor errors. Unfortunately, this

assumption has not been borne out experimentally (Burger & Wijnen, 1999; Howell & Vause, 1986; Howell, Williams, & Vause, 1987; Stromsta, 1986; Wingate, 1976) and clear evidence for a strong link between stuttering and phonological development is lacking (Melnick & Conture, 2000; Nippold, 2002), although further research is required (Conture, Zackheim, Anderson, & Pellowski, 2004; Melnick, Conture, & Ohde, 2003). A self-repair model that relies on sensorimotor error would be able to account for stuttering-like behavior observed in healthy speakers when exposed to delayed (Van Riper, 1982; Venkatagiri, 1980; Yates, 1963) or phase shifted (Stromsta, 1959) auditory feedback. These speakers were making repairs to perceived auditory error.

This view also suggests that errors in the fluent and dysfluent speech of PWS have a similar origin. There appears to be some empirical support for such a contention. Studies that have focused on the acoustic characteristics of sound/syllable repetitions indicate that F2 variations are observed (e.g., Agnello, 1975; Harrington, 1987; Yaruss & Conture, 1993) that are similar to the errors in the fluent speech of PWS. Furthermore, similar to the fluent speech of PWS, dysfluencies are most likely to occur on utterances with the vowel /i/. Analysis of Howell and Vause's (1986) data reveals that repetitions were more common for CVC tokens with /i/ as compared to other vowels. This information combined with the well known fact that most repetitions occur on syllables that start with consonants (Bloodstein & Ratner, 2008; S. F. Brown, 1938; Wingate, 1976) suggests that repetitions may be related to speech elements that require large formant transitions and are therefore most prone to auditory error.

The goal of modeling Experiment Two is to use the stuttering DIVA model to simulate a self-repair of sensorimotor errors by the monitoring subsystem and compare these simulations to the acoustic and kinematic characteristics of sound repetitions made by a single PWS.

2.4.1 Methods

We used the stuttering DIVA model to simulate productions of the nonsense phrase "a bad daba" (pronounced as /ebæd dæbə/)[8]. The *dysfluent production* started from the vocal tract configuration at rest. The DIVA model had to perform a rapid transition toward the initial configuration of the utterance (the configuration for the first vowel of "a bad daba"), and thus large errors were likely. The *fluent production* started from a vocal tract configuration that was closer to initial vowel configuration so that the initial transition was slow and large sensorimotor errors were unlikely.

To simulate repetitions, we implemented self-repairs in the DIVA model. In order to repair an error the monitoring subsystem disrupts the initial attempt (or the *1ˢᵗ attempt*) to produce the syllable by interrupting phonation as soon as possible, which takes 65ms given the DIVA model time constraints (Guenther et al., 2006)[9]. The CNS can indeed modify speech production that fast: same-trial compensation for formant shifting has been shown to require as little as 75ms (Tourville et al., 2008). Then, the monitoring

---

[8] The first vowel in "a bad daba" is typically produced as a schwa, but in accordance with the pronunciation of our subject, the DIVA model pronounces it as the vowel /e/.

[9] This estimation is based on the minimum time it takes to accelerate the tongue muscle, which includes 30ms from EMG to onset of acceleration (Guenther et al., 2006, p. 284). The laryngeal muscles, which are involved in interruption of phonation, have a similar delay relative to EMG, as demonstrated by Poletto et al. (2004, table 2).

subsystem initiates the next attempt (*2ⁿᵈ attempt*) to produce the syllable (the basic unit of articulatory execution in the DIVA model, see Guenther et al., 2006), but not before the feedback control subsystem issues corrective motor commands based on the magnitude and direction of the error. These commands reposition the articulators while voicing is turned off, so as to place them in an appropriate position for the upcoming attempt to produce the syllable correctly. If the 2ⁿᵈ attempt still contains an auditory error, the monitoring subsystem will initiate another self-repair. The cycle repeats until speech output is either error-free or the error is small enough to be handled by the sensory feedback loops. At this point no further disruption occurs. Each repetition consists therefore of one or more disrupted attempts followed by a *final complete production* of the utterance. This account of error repair is drawn from previous theoretical work (Postma & Kolk, 1993; Stromsta, 1986; Van Riper, 1982).

The self-repair simulations were compared with kinematic and acoustic data based on a single speaker drawn from the Walter Reed-Western Michigan University Stuttering Database, a large speech acoustic and physiological dataset of adults who do and do not stutter (McClean et al., 2004; Tasko, McClean, & Runyan, 2007). The speaker was 19 year old male (denoted CXX) with no reported history of stuttering therapy. A behavioral analysis of videotaped samples of CXX's monologue and oral reading was performed by expert clinicians. His score on the Stuttering Severity Instrument-3 (Riley, 1994) was 26, which placed him in the range of a moderate fluency impairment. All analysis reported here is based on CXX's repeated production of a nonsense phrase "a bad daba" (/ebæd dæbə/) performed at habitual rate and loudness, and pausing at a

comfortable rate between each token. This speaker was selected because he consistently repeated the initial vowel of the test phrase.

The participant was seated in a sound-treated room while chest wall motion, orofacial movement and speech acoustics were recorded. Recordings were obtained of the two-dimensional positions of the lips, tongue blade, jaw, and nose within the midsagittal plane by means of a Carstens AG100 Articulograph (Carstens Medizinelektronik GmbH, Lenglern, Germany), an electromagnetic movement analysis system commonly known as EMA. Three sensor coils (3 mm x 2 mm x 2 mm) were attached with biomedical tape to the bridge of the nose and the vermilion borders of the upper (UL) and lower lip (LL). Two more sensor coils were attached with surgical adhesive (Isodent) to the tongue blade (TB) approximately 1 cm from the tip and at the base of the mandibular incisor (MI). Sensor locations are schematically presented in Fig. 2-3(c). The acoustic speech signal was transduced with a Shure M93 miniature condenser microphone (Shure, Inc., Niles, IL) positioned 7.5 cm from the mouth and the microphone-amplifier setup was calibrated to permit measurement of absolute sound pressure levels. The orofacial kinematic and acoustic signals were digitized to a computer at 250 Hz per channel and 16 kHz respectively. The subject performed approximately 12-15 speech tasks, each 30 seconds in length.

After data acquisition, the lip, jaw, and tongue movement signals were low-pass filtered (zero phase distortion fifth-order Butterworth filter) at 8 Hz and the nose signal at 3 Hz. While head movements during recording were slight (<1.0 mm), nose sensor movements were subtracted from the lip, tongue, and jaw movement signals in the X and

Y dimensions in order to minimize any head movement contributions. The kinematic signal was downsampled to 146Hz and the acoustic signal was upsampled to 22.050 kHz to fit the file format supported by the TF32 (Time-Frequency analysis for 32-bit windows) software (Copyright 2000 Paul Milenkovic. Revised May 7, 2004). The TF32 software was then used for data analysis and presentation. To compare the data to the model simulations, we calculated the approximate positions of the pellets on DIVA's artificial vocal tract (Maeda, 1990) and sampled their X and Y coordinates during the simulations. These coordinates were then scaled to enable rough comparison with individual subject data. Acoustic analysis was carried using the SpeechStation 2 software (Copyright 1997-2000 Semantic Corporation. Version 1.1.2). Formants were identified by the first author using linear predictive coding (LPC) technique (12 coefficients) in combination with a wideband spectrogram. Measurements were taken in the first glottal pulse of each disrupted attempt or fluent production with the condition that both the first and second formants be visible on the spectrogram.

2.4.2 Results

Fig. 2-3(a) is an F1-F2 chart showing the formant values of both fluent and dysfluent productions of the initial vowel in the phrase "a bad daba" (/ebæd dæbə/) made by the stuttering DIVA version (in blue) and the test subject CXX (in other colors). As expected, a self-repair was triggered in the beginning of the dysfluent production by the DIVA model. The 1st attempt to produce the /e/ is marked by the big square (its location specifies the F1 and F2 of the auditory feedback at the beginning of the attempt), which is

connected by a dashed line to the small square, marking the 2nd attempt to produce this sound. Another line connects that square to a circle that marks the final, complete production of /e/. Four separate sound repetitions sequences made by CXX are depicted in the same manner, with each repetition plotted in a different color. A blue star marks the F1-F2 position of the initial vowel for the DIVA model's fluent production of the utterance. A black star similarly marks the F1-F2 positions for CXX's fluent production. The DIVA model and CXX used slightly different frequencies to produce /e/ because they have different vocal tract dimensions, a factor that modulates the formant values produced by a speaker (Peterson. & Barney, 1952).

The F2 values of the 1st and 2nd attempts by the DIVA model are 1730Hz and 1810Hz, respectively. The final complete production has an F2 of 1920Hz. Second formant values for all repetitions made by CXX (those that are plotted in Fig. 2-3(a), as well as seven additional repetitions), are listed in Table 2-1. For each repetition the F2 values at the starting points of the disrupted attempt(s), as well as of the final complete production, are listed. An unpaired samples t-test ($p < 0.0004$) showed that there was an increase in F2 from the disrupted attempts to the final complete production. Also listed in the table are the durations of CXX's disrupted attempts. The durations of the 1st and 2nd attempts in the DIVA model simulation were 66ms and 76ms respectively, indicating that the errors were detected shortly after voicing onset (1ms and 11ms respectively).

(a)

(b)

(c)
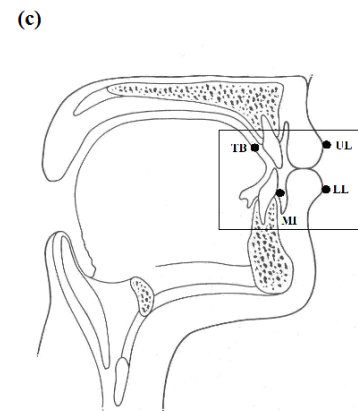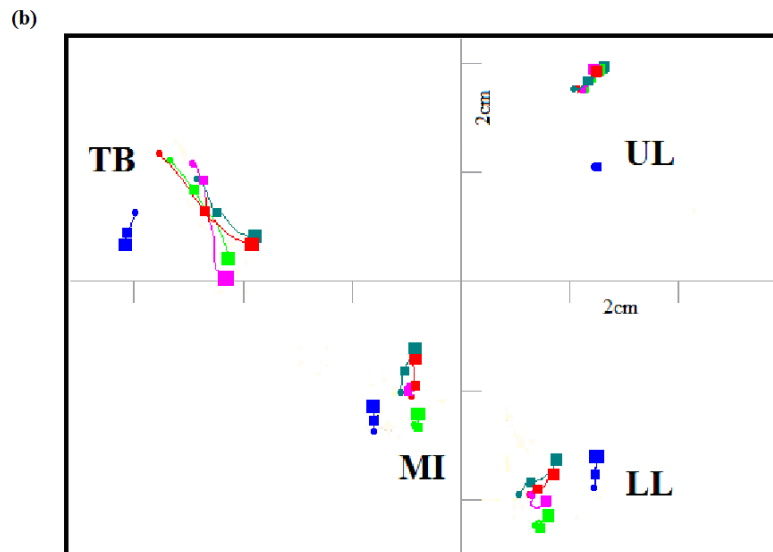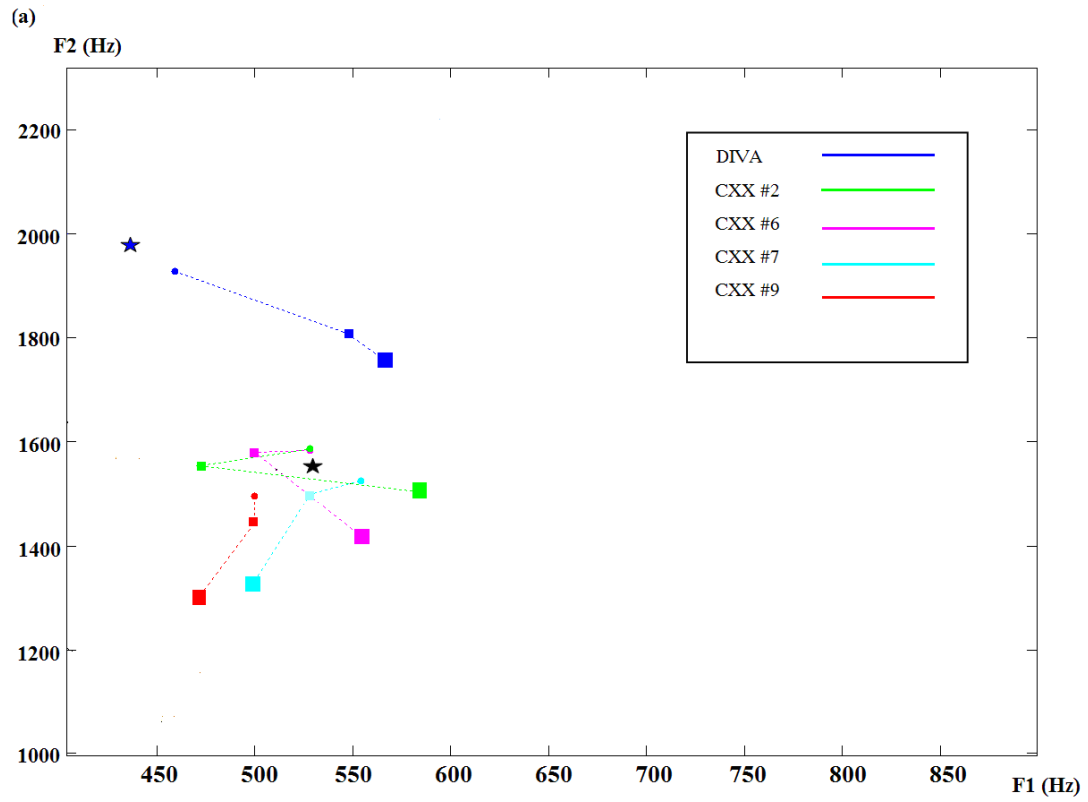
Fig. 2-3. One simulated (blue) vs. four recorded (other colors) repetitions on the first vowel (pronounced as /e/) in "a bad daba". Simulations were performed with the stuttering version of the DIVA model. Recordings are of the stuttering speaker CXX. For each repetition, big and small squares mark the beginning points of the 1st and 2nd attempts respectively. Circles mark the beginning points of the final complete production that follows the repetition. (a) Acoustic data. Circles and squares indicate formant values on a F1/F2 chart. A dashed line connects the markers of each repetition. The serial numbers of the repetitions listed in Table 2-1 are specified in the legend. Stars mark the formants of the fluent productions of /e/ (blue for the simulation and black for CXX). (b) Articulatory data. For each repetition, the trajectories (X and Y coordinates) of the 4 pellets used for data acquisition are plotted. UL – upper lip. LL – lower lip. MI – Mandibular incisor. TB – tongue blade. The trajectories follow pellet locations starting at the beginning of the 1st attempt and ending at the beginning of the final complete production. (c) A schematic diagram (subject identity unknown) showing pellet approximate locations on a Sagittal plane of the vocal tract.

Table 2-1. F2 values during repetitions made on the first vowel (pronounced as /e/) of "a bad daba" by the stuttering speaker CXX.

#	1st disrupted attempt		2nd disrupted attempt		Final complete production
	F2 (Hz)	Duration (ms)	F2 (Hz)	Duration	F2 (Hz)
1	1465	130	…	…	1709
2*	1506	80	1555	120	1588
3	1465	130	…	…	1627
4	1465	130	…	…	1627
5	1546	60	1587	70	1543
6*	1415	180	1581	20	1584
7*	1329	90	1498	120	1526
8	1302	10	1546	140	1627
9*	1303	100	1448	110	1494
10	1383	70	…	…	1505
11	1343	120	…	…	1546
Mean	1411 (85)	100 (45)	1535	97 (44)	1580 (65)

* repetition is also plotted in Fig. 2-3.

Fig. 2-3(b) shows the articulatory movements for the repetitions made by the DIVA model and CXX. For each repetition (depicted in the same color as in Fig. 2-3(a)), four continuous lines trace the positions of the 4 pellets (TB, UL, MI, and LL) from the beginning of the first attempt (big squares) to that of the second attempt (small squares), and from there to the beginning of the final complete production (circles). While the simulation generally agrees with CXX data regarding the direction to which the articulators are moving, the tongue blade of DIVA does not seem to move backwards similarly to CXX' tongue blade (even though in both cases the tongue blade moves upwards). This difference in orientation is not that surprising given that there are differences in vocal tract morphology (and, correspondingly, differences between the relationship between articulator positions and formant values) between CXX and the simplified Maeda (1990) articulatory synthesizer used in the DIVA simulations as well as the details of the reference frame used for each.

2.4.3 Discussion

The speaker data shows that within a single repetition, the disrupted attempts show a formant pattern that converges toward the F2 value for the fluent production of the vowel. This suggests there may be some corrective action occurring through the repetition. The DIVA simulation, which is based upon the self-repair hypothesis, also shows such a pattern suggesting that it is plausible that repetitions may result from a self-repair of sensorimotor error. It is also noteworthy that the self-repair investigated here primarily affects F2. This can be explained by the second formant being the primary locus of error in stuttering (Section 2.3.3).

Patterns similar to our empirical and simulation results have been previously observed. Stromsta (1986) reported on a child who repeated part of the initial vowel in the word "apple." The disrupted attempts did not show any evidence of acoustic transitions one would expect for a vowel followed by a /p/ suggesting an error. In line with our simulation, the error was limited to the disrupted attempts, and was resolved when the utterance was finally produced. A variety of studies have described F2 abnormalities during repetitions. In Howell and Vause's (1986) data for repetitions on /CiC/ syllables, the F2 values were abnormally low in the disrupted attempts of the repetitions, but shifted to higher values in the final, complete production. Harrington (1987) demonstrated repetitions where abnormal F2 transitions were present in the disrupted attempts but not in the final complete production. Similar abnormal F2 transitions were also observed in 16% to 45% of the disrupted attempts in the Yaruss and Conture (1993) study of repetitions by children who stutter. Unfortunately, the aforementioned findings cannot be directly compared to our DIVA simulation since none of the studies could generalize the abnormalities to all repetitions (see also Subramanian et al., 2003), for one of several reasons: (a) because relatively few repetitions were analyzed (Howell & Vause, 1986; Yaruss & Conture, 1993); (b) because no quantitative measures were used (Harrington, 1987); or (c) because each participant repeated on different words (Yaruss & Conture, 1993).

The acoustic data cannot reveal the speech motor behavior during the silent periods between the disrupted attempts of the repetition. Therefore, we also evaluated articulatory kinematic data. Fig. 2-3(b) shows that CXX repositions the articulators

between the disrupted attempts. Between the 1st and 2nd attempts, and between the 2nd attempt and the final complete production, the tongue blade moves upwards and does not return to its initial position. The same pattern of repositioning is also observed in the repetition made by the stuttering DIVA version. This suggests that both the speaker and the DIVA model may be repositioning articulators by issuing corrective motor commands; commands that are based on the direction and extent of the error that triggered the self-repair. These results are in line with reports by Zimmermann (1980a) who reported "consistent repositioning of articulators occurs during oscillatory behaviors [i.e., repetitions]…" (p. 117).

The acoustic data in Table 2-1 and Fig. 2-3(a) suggest that the repositioning of articulators by CXX affected the formants in a consistent way. This can be best observed when CXX made 2 attempts before producing the utterance fluently. In almost all of these repetitions (#2, #6-#9) there was a gradual increase in F2, from the 1st to the 2nd attempt, and from there to the final complete production. The same gradual increase in F2 was observed during the DIVA model simulated repetition (from 1730Hz to 1810Hz to 1920Hz). Moreover, the constant increase in F2 of the auditory feedback seems to bring it toward the F2 of the fluent production (1975Hz) where error, if exist, is small. This combined observation of acoustics and kinematics supports the view that articulatory repositioning during repetitions brings the auditory feedback closer to the desired auditory target[10].

---

[10] Based on kinematic data only, Zimmerman (1980b) reached a different conclusion: that the repositioning serves to alter the inappropriate muscle activity that "had a disruptive effect, via reflex pathway, on the muscles necessary for fluent production" (p. 131).

Postma and Kolk (1997) rejected the theory that motor errors lead to moments of stuttering, in part because of their assumption that auditory feedback is too slow to allow fast detection of such errors. Our simulation suggests that auditory feedback is fast enough to account for the time it takes to trigger repetitions: even the longer of the two disrupted attempts by the stuttering DIVA version, an attempt that lasted 76ms, was fast enough to account for the disrupted attempts by CXX that were 99ms long on the average (Table 2-1) as well as durations of disrupted attempts reported in the literature, which range anywhere from 80-150 ms (Stromsta, 1986; Yaruss & Conture, 1993).

The relatively rapid disruptions due to sensorimotor errors, often in the middle of the syllable, justify the association between such errors and sound/syllable repetitions. Moreover, PNS who seldom make such errors should make fewer part-syllable repetitions than PWS, not only in absolute terms, but also in proportion to the total number of repetitions (part-syllable, whole-syllable, and multiple-syllable). This is indeed the case, since not more than 12%-18% of the repetitions made by PNS are of part-syllables (Levelt, 1983, p. 61; McClay & Osgood, 1959), in comparison to 46%-63% of the repetitions made by PWS (Sheehan, 1974, p. 202; Soderberg, 1967).

2.5   Modeling experiment 3: Slowed/prolonged speech reduces the frequency of sound/syllable repetitions

It has been frequently reported that PWS stutter less when speaking at a slower rate (for review see Andrews et al., 1983, p. 232; Starkweather, 1987, p. 192; Van Riper, 1982, p. 401), or prolong their words (Davidow, Bothe, Andreatta, & Ye, 2008; Ingham et al., 2001); two conditions that may reduce the speed in which PWS move their articulators.

We hypothesize that slower articulatory movements may reduce the extent and frequency of errors made by PWS (Max et al., 2004**,** p. 114); as a result, fewer repetitions are triggered. This hypothesis can explain, for example, why PWS who speak slower (after fluency-shaping treatment) have less vowel centralization (hence, less auditory errors) compared to PWS that speak at a normal rate (Blomgren et al., 1998). The current simulation was designed to demonstrate that the frequency of errors, and as a consequence also repetitions, drops significantly when the DIVA model is required to move the articulators at a rate that is 50% of the normal speaking rate.

2.5.1 Methods

Before simulating slowed/prolonged speech, we first verified that the DIVA model can account for the difference in the frequency of repetitions made by PWS and PNS. To that end, we simulated both the stuttering and the non-stuttering versions of DIVA producing "good dog" at a normal speech rate (*normal speed condition*). Then, to permit the stuttering DIVA version to use slower movements, the target region of the utterance "good dog" was linearly stretched in time so as to double its duration (*slow speed condition*). The stretching in time reduces the slope of formant transitions. Thus, the correct production of the utterance does not require as fast articulatory movements as before. This manipulation also makes the steady-state portion of vowels longer and increases the pause between words. To analyze how the repetitions distribute over the phonemes of the utterance, in each one of the simulations, the DIVA model produced the utterance 500 times.

2.5.2 Results

Fig. 2-4 shows the simulation results for the stuttering DIVA version in the normal speed condition (plot a) and the slow speed condition (plot b), and that of the non-stuttering DIVA version at normal speed (plot c). For each simulation, the auditory feedback during one of the fluent productions of the utterance is displayed in the top of the corresponding plot. The distribution of repetitions over the phonemes of the utterance is given by the frequency histogram in the bottom of the plot.
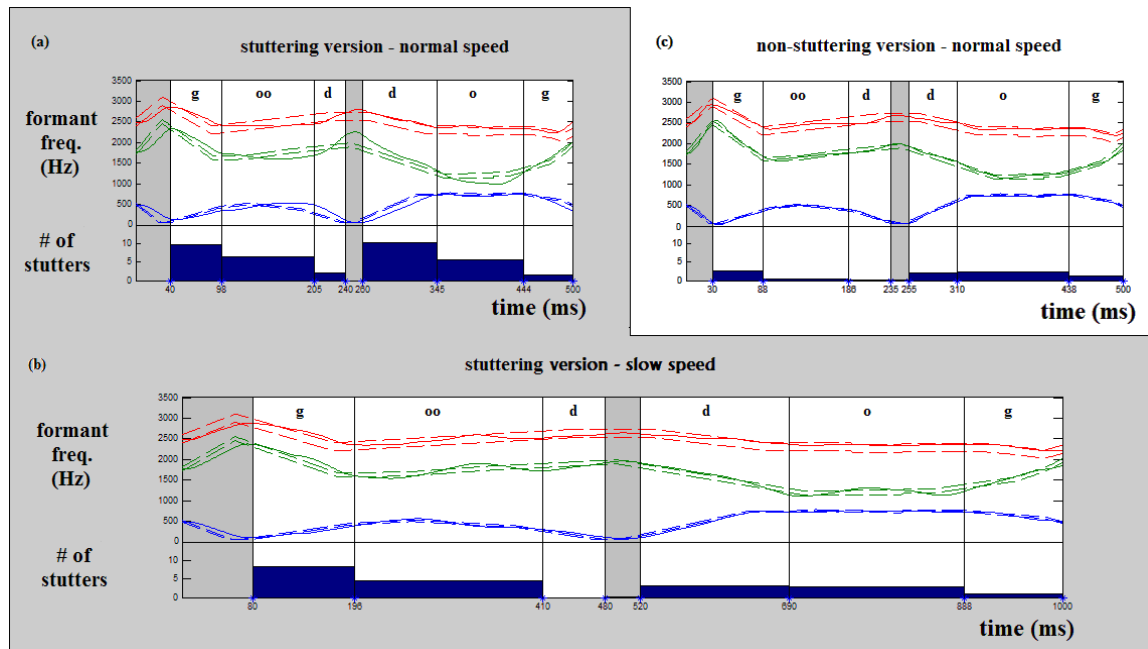
Fig. 2-4. Acoustics and frequency of sound/syllable repetitions made by the stuttering version of the DIVA model at normal (a) and slow speeds (b), and repetitions made by the non-stuttering version of the DIVA model, at normal speed only (c). The top of the plots use the same conventions as in Fig. 2-2. In the bottom of the plots, histograms show how many repetitions occurred on each phoneme of the utterance in 500 productions of "good dog" (silent periods are shaded).

Consistent with modeling Experiment One, at the normal speed condition the auditory errors (mismatch between auditory feedback and target region) of the stuttering

DIVA version were larger than those of the non-stuttering version. The difference in the auditory error size resulted in a much greater frequency of repetitions for the stuttering (43/500) vs. the non-stuttering DIVA model (9/500). The stuttering DIVA version differs from the non-stuttering version also in having the greatest number of repetitions on word-initial positions. In the slow speed condition the stuttering DIVA version had much smaller errors and fewer repetitions (21/500) than in the normal speed condition.

2.5.3 Discussion

The simulation confirmed that in the slow speed condition, reducing movement speed by half reduced the size of the auditory errors, thus also reducing the frequency of repetitions. According to the simulation, the reduction was an outcome of the fact that the target region had slower formant transitions, while the system time lags remained unchanged. Given the same feedback control delays, the flatter slopes of the target region transitions were much easier for the stuttering DIVA version to follow. The agreement of the simulation with the reduction in stuttering observed when PWS speak half as fast — when instructed to speak slower (Andrews, Howie, Dozsa, & Guitar, 1982; Ingham, Martin, & Kuhl, 1974), or when instructed to prolong their speech (Perkins, Bell, Johnson, & Stocks, 1979) — provides further evidence for the applicability of our theory to PWS. However while the effect of slowed/prolonged speech generally agrees with our simulations, it is difficult to conduct a quantitative comparison due to the paucity of the data for these procedures in a controlled setting (Ingham et al., 2001, p. 1230). Specifically, no data are available regarding the fraction of sound/syllable repetitions out of the total reduction in stutterers.

Further complications arise from the way PWS interpret the instruction to slow down (Andrews et al., 1982; Healey & Adams, 1981) or prolong (Ingham et al., 2001, p. 1230) their speech. The difficulty is more acute for slowed speech where PWS prefer prolonging pauses between words to prolonging the words themselves (Healey & Adams, 1981). For example, in one reduced speech rate study (Andrews et al., 1982), all three PWS at least doubled the duration of pauses, while only one of the PWS significantly increased the duration of words. In order to overcome this problem, many currently popular stuttering treatment programs train PWS to stretch sounds and not pauses (Curlee, 1999); a method that has the effect of reducing the speed and increasing the duration of articulatory movements associated with transitions (Tasko et al., 2007). Ingham et al. (2001), for example, developed the MPI (Modifying Phonation Intervals) treatment program which trains PWS to specifically reduce the frequency of short phonation intervals, thus encouraging prolongation of words. The results of this treatment are encouraging — all of the participants in the program achieved stutter-free and natural-sounding speech.

2.6 Modeling Experiment 4: Masking noise reduces the frequency of sound/syllable repetitions

Masking noise (constant binaural white noise) significantly reduces the average frequency of stuttering (for review, see Andrews et al., 1983; Bloodstein, 1995, p. 345; Van Riper, 1982, p. 380), with the reduction being the greatest for sound/syllable repetitions (Altrows & Bryden, 1977; Conture & Brayton, 1975; Hutchinson & Norris, 1977). Some have argued that masking noise completely blocks auditory feedback

(Sherrard, 1975; Stromsta, 1972; Van Riper, 1982, p. 382), but this is inconsistent with PWS' frequent reports that they keep hearing themselves above the noise (Adams & Moore, 1972; Shane, 1955). We suggest, instead, that masking noise reduces the quality of auditory feedback (cf. Garber & Martin, 1977, p. 239; Starkweather, 1987, p. 184) by decreasing the signal-to-noise (SNR) ratio (Liu & Kewley-Port, 2004). The decreased SNR may deteriorate the precision with which PWS perceive formant values, thus preventing them from detecting small errors (since error detection requires the speaker to perceive small mismatches between the formants of the auditory feedback and those of the target region; see Section 2.3.2). Not triggering repetitions to repair small errors, PWS will then stutter less. Our treatment of masking noise is supported by a study from Postma and Kolk (1992) reporting that noise caused PWS to detect a smaller fraction of the speech errors they made while producing tongue-twisters. Relative to the number of errors, PWS that spoke in noise also had fewer dysfluencies (repetitions among them).

The current account makes predictions regarding changes in masking noise intensity as well. Since louder noise means a lower signal-to-noise ratio (and detection of fewer errors), the frequency of repetitions should decrease as noise rises. Indeed, in a series of experiments that modulated masking noise intensity (Adams & Hutchinson, 1974; Cherry & Sayers, 1956; Maraist & Hutton, 1957; Shane, 1955), it was found that the frequency of stutters is reduced to a greater extent under louder noises. Here we investigate whether the stuttering version of the DIVA model makes fewer repetitions as it is exposed to higher levels of masking noise.

2.6.1 Methods

Liu and Kewley-Port (2004, p. 3125) demonstrated that masking noise increases the threshold for discriminating between two perceived formants. Similarly, we assume that masking noise increases the threshold for detection of an auditory error (the threshold for detecting the mismatch between the auditory feedback formants and those of the target region). In light of Liu and Kewley-Port's report of a close-to-linear relation between masking noise and formant discrimination threshold, we use the following formula to calculate the threshold for error detection, $T_{detection}$, for a given masking noise intensity, $I_{masking}$:

$$T_{detection}\left(I_{masking}\right) = I_{masking} * \xi$$

where the constant ξ is set to 2, $T_{detection}$ is measured in Hz, and $I_{masking}$ is measured in dB.

We conducted 10 simulations of the DIVA model repeating the utterance "good dog", with each simulation having different noise intensity $I_{masking}$ ranging from 0 dB to 90 dB in steps of 10 dB. In each simulation, the DIVA model could only detect auditory errors that were greater than $T_{detection}\left(I_{masking}\right)$ Hz. Moreover, the errors that could be detected were perceived as smaller than they really were; an error that was perceived as X Hz under normal conditions was perceived under masking noise as only $X - T_{detection}\left(I_{masking}\right)$ Hz in size.

2.6.2 Results

Fig. 2-5 shows, for each noise intensity $I_{masking}$, the number of repetitions made by the stuttering DIVA version (dashed line). The thresholds of error detection for masking

noise intensities ranging from 0 to 90 dB are depicted at the bottom of Fig. 2-5. In silence (0 dB), the DIVA model made on the average 3.7 repetitions per 100 syllables; the model made fewer repetitions at higher intensities, with only 1.1 repetitions on average with a masking noise of 90 dB.
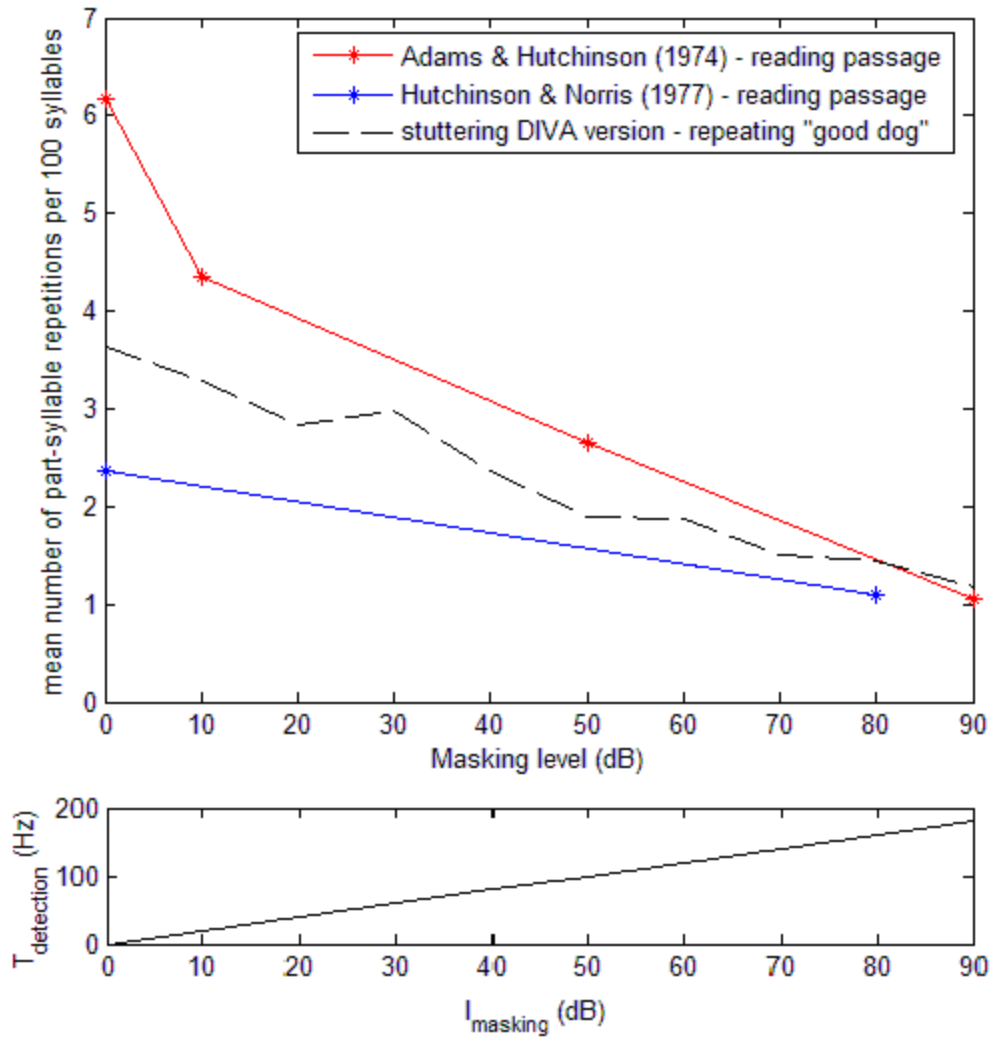
Fig. 2-5. Frequency of sound/syllable repetitions made by the stuttering version of the DIVA model and selected studies under masking noise of different intensities. The number of repetitions (upper y-axis) under different noise intensities (x-axis) contrasts the DIVA model simulation and the experimental results. The threshold for detecting errors, $T_{detection}$ (lower y-axis), is indicated by the black line. Study results are expressed as repetitions per 100 syllables (see Section 2.6.3 for details).

2.6.3 Discussion

The simulations confirm that the DIVA model reduces the frequency of repetitions under masking noise and that the reduction is greater for louder noise. Moreover, as can be seen in Fig. 2-5, the simulation results resemble the estimated reduction in repetitions in the data reported by Adams and Hutchinson (1974)[11]. In that study, the estimated number of repetitions in the no-noise condition was much higher than in the 10 dB condition because headphones were only worn in the noise conditions. Speaking with headphones is a novel situation that reduces stuttering in its own right (e.g., T. Brown, Sambrooks, & MacCulloch, 1975). The simulations are also consistent with the results of a study that directly measured the frequency of repetitions in silence and at 80 dB (Hutchinson & Norris, 1977). Other studies have reported results that are similar to the simulations (Conture & Brayton, 1975; Maraist & Hutton, 1957). In conclusion, the agreement between the simulations and the experimental data suggests that the effect of masking noise on PWS is similar to its effect on the stuttering DIVA version: the masking noise lowers the signal-to-noise ratio, which prevents detection of errors and subsequently reduces the frequency of repetitions.

To demonstrate that masking noise acts on the signal-to-noise ratio, repetitions should be also affected by modulation of the "signal" part of the ratio — the auditory feedback (Starkweather, 1987**,** p. 184, 239). This was indeed confirmed by Garber and

---

[11] Adams and Hutchinson (1974) reported the number of all stutters combined. Therefore, the estimation of the frequency of sound/syllable repetitions was based on the fraction of part-syllable repetitions out of the total number of stutters reported by Hutchinson and Norris (1977), a fraction that seems to vary with noise intensity. We disregarded whole-syllable repetitions because unlike part-syllable repetitions, their frequency barely decreased in the noise condition.

Martin (1977): on the average PWS had less stutters (repetitions among them) with lower auditory feedback (using low instead of high vocal intensity, masking noise intensity being equal). Another explanation is possible however: the PWS had to keep their vocal intensity low despite the noise, thus resisting the Lombard effect. Such an intentional change in speech style may reduce stuttering on its own right (Alm, 2004; Bloodstein & Ratner, 2008).

It is noteworthy that not all stuttering in noise studies can be accounted for by a reduced signal-to-noise ratio view. The suggestion that noise reduces stuttering through distraction (Bloodstein, 1995, p. 350) can explain the effect of some noise forms that do not significantly reduce the signal-to-noise ratio: noise that is played for less than 20% of the time (Murray, 1969); noise that is played monaurally (Yairi, 1976); narrow-band noise played at 4 or 6Khz, much higher than the range of the first 3 formants (Barr & Carmel, 1969); or pure tones of various frequencies (Cherry & Sayers, 1956; Parker & Christopherson, 1963; Saltuklaroglu & Kalinowski, 2006; Stephen & Haggard, 1980) [12]. Yet, masking noise acts on repetitions by more than just distracting PWS because for a given intensity it is more effective than most other noise forms (Murray, 1969; Yairi, 1976). Moreover, the effectiveness of masking noise after prolonged use (e.g., Altrows & Bryden, 1977; MacCulloch, Eaton, & Long, 1970) cannot be explained if masking noise is only a distractor, which presumably loses its effect with time.

---

[12] Binaural white noise that is played only during silent periods (Sutton & Chase, 1961; Webster & Dorman, 1970) was not included in the list because albeit not overlapping in time with phonation, it may still reduce the signal-to-noise ratio in the initial tens of milliseconds of words (where errors are most likely to occur, see Section 2.7) via forward masking (Moore, 1995). Similarly, bimanual white noise that is turned on shortly after a word begins and up to its end (see Chase & Sutton, 1968) may mask the initial portion of the word via backward masking (Moore, 1995).

An underlying hypothesis guiding the simulations is that masking noise eliminates repetitions by preventing PWS from detecting errors rather than by making PWS move their articulators differently. In keeping with this, masking noise was not found to significantly change the vowel formant values produced by PWS (Klich & May, 1982) or the percentage of speech errors (including distorted vowels) they make on a given utterance (Postma & Kolk, 1992). Since masking noise unexpectedly barely reduced stuttering in Klich and May's study, their results should be treated with caution, but the same argument cannot be used to reject Postma and Kolk's study, where a significant reduction in dysfluencies was observed in the masking noise condition.

2.7 General discussion and Conclusions

Based on recent insights into the neural control of speech production gained from neurocomputational modeling, we presented a specific hypothesis about a possible source of stuttering. The hypothesis suggests that due to the time lags inherent in feedback control, over-reliance on auditory feedback (a consequence of impaired feedforward commands) causes mismatches between the desired speech output and actual auditory feedback, especially in rapid formant transitions. When the auditory error becomes too large, it is detected by the CNS which tries to self-repair the error by immediate disruption of phonation, repositioning of the articulators, and an attempt to reproduce the erroneous syllable. The repeated disrupted attempts to produce the syllable, until it is produced without errors, are what clinicians formally label sound/syllable repetition.

The hypothesized speech motor control abnormalities are supported by simulations that replicated published work on fluent speech of PWS. Furthermore,

simulations of self-repairs that are being initiated in order to eliminate these abnormalities resembled our own data on PWS dysfluent speech. We also surveyed experimental results that reject most of Postma and Kolk's (1993**,** p. 482) objections to speech motor control theories of stuttering that assume sensorimotor errors: errors and stutters are more than merely co-occurring phenomena because the probability for a stutter depends on error size (Section 2.2.2), and errors are not a secondary consequence of a lifetime stuttering (maybe due to some sort of compensation strategy; see Armson & Kalinowski, 1994), because they are evident already in young age (Sections 2.3.3 and 2.4.3, cf. Conture, 1991). Not enough data were collected however to reject the possibility that the apparent errors are due to a gross change in the speech apparatus. It was proposed that PWS actively alter the length of their vocal tract (Prosek et al., 1987) or it may change in size due to abnormal development.

The design of the DIVA model emphasizes the importance of the role of auditory information in speech in monitoring for errors (monitoring subsystem) and in feedback-based motor control (feedback control subsystem). Based on experimental results we assumed that masking noise heavily interferes with the monitoring for errors (thus, significantly reduces the number of repairs), but not with the feedback-based motor control (thus, does not significantly change actual speech production); whether feedback-based motor control is spared due to the somatosensory information that it utilizes in addition to the auditory information is a matter of future investigation. The multiple roles of auditory information may help explain also other fluency enhancers that perturb or interfere with auditory feedback, such as DAF (Delayed Auditory Feedback), FAF

(Frequency-shifted Auditory Feedback), and chorus speech (Andrews et al., 1983; Bloodstein, 1995; Macleod, Kalinowski, Stuart, & Armson, 1995). Hearing loss is another condition that affects auditory feedback, and according to the current account, should interfere with monitoring for errors as well. Indeed, the incidence of stuttering in the hearing impaired is low (Montgomery & Fitch, 1988; Starkweather, 1987**,** p. 243; Van Riper, 1982**,** p. 47; Wingate, 1976**,** p. 216). That no sound/syllable repetitions were observed in the only completely deaf child who stuttered in the Montgomery and Fitch survey (1988, p. 1933), further supports the view that repetitions of this kind are especially prone to low quality auditory feedback (see Section 2.6).

In addition to explaining certain conditions that alleviate stuttering, over-reliance on auditory feedback can also explain why stutters are most likely to occur during (or immediately after) the production of phonemes in word-initial position (Andrews et al., 1983; Bloodstein, 1995; Wingate, 2002**,** p. 327). We argue that because auditory feedback is not available in the silent periods of speech, PWS must depend then on the impaired feedforward commands that are not capable of bringing the articulators to their correct positions[13]. This will lead to big auditory error at voicing onset and increased likelihood of a repetition. Modeling experiments One and Three support this argument by showing that the stuttering DIVA version makes the biggest auditory errors on word-initial phonemes; that these phonemes are also the most likely to be stuttered was confirmed by modeling experiment three.

---

[13] While PWS may still use the somatosensory feedback loop in silent periods, it probably cannot compensate for the total absence of auditory feedback. The somatosensory feedback is not helpful, for example, when PWS are deprived of visual feedback in a non-acoustic oral task (Loucks & De Nil, 2006b).

The correspondence of the DIVA model components to anatomical locations in the brain (Guenther et al., 2006; Tourville et al., 2008) provides us with an additional way to test our hypothesis: by comparing the activities of the model components to activities of brain regions measured using functional brain imaging during the fluent and dysfluent speech of PWS (for use of this process of comparing model simulations and experimental results for the speech of PNS, see Guenther, 2006; Guenther et al., 2006; Tourville et al., 2008). Based on our recent association of the motor mechanisms for error correction with right hemisphere inferior frontal cortex (Tourville et al., 2008), we can account for the over-activation of that region often reported in the speech production of PWS as compared to PNS (S. Brown, Ingham, Ingham, Laird, & Fox, 2005). The association of auditory error cells with the posterior superior temporal gyrus (Tourville et al., 2008) predicts PWS to have differential activation in that region as well, but further model refinements and simulations are needed to ensure that the sign and size of the model's activity are comparable with physiological measurements. Lastly, mapping the anatomical location of the components of the newly defined monitoring subsystem may be possible when comparing model activities to imaging data collected during sound/syllable repetitions. Since this is the subsystem that initiates repetitions, its components are expected to be extremely active in these instances.

3 BASAL GANGLIA DYSFUNCTION MAY LEAD TO DYSFLUENCIES

3.1 Introduction

Despite the fact that approximately 1% of the population stutters (Van Riper, 1982), the etiology of the disorder is still unknown. The primary goal of this chapter is to enhance our understanding of the neurological basis of stuttering. This problem is approached here using computational modeling to provide mechanistic accounts for data in the literature and generate testable hypotheses to guide future experiments. This approach seeks to unite the examination of brain and behavior in speech production, continuing in the spirit of recent instantiations of the DIVA (Directions into Velocities of Articulators) and GODIVA (Gradient Order DIVA) models (Bohland et al., in press; Guenther, 1995, 2006; Guenther et al., 2006; Guenther et al., 1998; Nieto-Castanon et al., 2005; Tourville et al., 2008).

Over the last two decades, brain imaging studies of stuttering have been advancing our understanding of the disorder. Of particular interest is the recent discovery of abnormalities in the brains of persons who stutter (PWS). In this chapter we discuss two of these abnormalities: a structural abnormality in white matter fibers (WMF) beneath the left precentral gyrus (Chang et al., 2008; Sommer et al., 2002; Watkins et al., 2008), and evidence of elevated dopamine (DA) levels (Wu et al., 1997), possibly caused by increased dopamine release from the substantia nigra pars compacta (Watkins et al., 2008). Although both abnormalities have been suggested as possible causes of stuttering (e.g., Alm, 2004; Giraud et al., 2007), it is unclear what functional neural circuit is implicated in the fiber bundle impairment, or what brain regions are affected by the

elevated dopamine levels. Also unknown are the functional consequences of these abnormalities, and how they lead to moments of stuttering.

We hypothesize that the structural abnormalities are located in those corticostriatal projections that carry corollary discharge of motor commands sent from the ventral premotor cortex (vPMC) to the ventral primary motor cortex (vMC), and that the resulting transmission errors prevent the basal ganglia (BG) from detecting the context for initiating the motor program for the next syllable (cf. Alm, 2004, p. 358). Concerning the second abnormality, it is our claim that elevated dopamine levels affect the putamen (part of the striatum), and hamper the BG's ability to bias cortical competition (among motor programs for similar syllables) in favor of the motor program appropriate for the next syllable.

The BG, which are involved in speech production (Crosson, 1992; Guenther, 2008), were first linked to stuttering only indirectly: the association was made due to the BG's interconnection with the supplementary motor area (SMA), which may be involved in the disorder as well (Caruso, Abbs, & Gracco, 1988). At around the same time, it was shown that there is a significantly higher incidence of BG involvement in brain lesions that cause speech dysfluencies than in those that do not (Ludlow, Rosenberg, Salazar, Grafman, & Smutok, 1987). Moreover, it has been noted that the neurogenic stuttering associated with BG lesions is similar to developmental stuttering in various behavioral and clinical dimensions (Heuer, Sataloff, Mandel, & Travers, 1996; Koller, 1983).

The strongest evidence for the involvement of the BG in stuttering, however, comes from drug studies. In repeated findings, drugs that block dopamine D2 receptors

have been shown to be effective in reducing stuttering (Brady, 1991; Maguire, Yu, Franklin, & Riley, 2004; Stager et al., 2005). It is likely that these drugs act on the BG because the BG are major targets of dopamine in the brain. D2 antagonists have been developed for the treatment of schizophrenia and bipolar disorder, and their use in the treatment of stuttering was motivated by the similarity between stuttering and Tourette's syndrome (another disorder treated with D2 antagonists): both begin in childhood, and both occur more frequently in males than in females. Unfortunately, clinical use of most D2 antagonists is problematic due to side effects. Risperidone may cause hyperprolactinemia, leading to sexual dysfunction; olanzapine may cause weight gain and increased risk for diabetes (Tran, Maguire, Franklin, & Riley, 2008); ziprasidone is associated with heart problems (Maguire & Snyder, 2008, August 18); and haloperidol, a first-generation dopamine antagonist, can cause tardive dyskinesia (Ludlow & Loucks, 2003). Recently-developed partial dopamine agonists (or dopamine system stabilizers, see Stahl, 2001), which are similar to D2 antagonists in their action, have fewer side effects, and are therefore more suitable for long-term administration (Tran et al., 2008).

The BG's involvement in stuttering is also supported by functional imaging studies (Braun et al., 1997; S. Brown et al., 2005; Giraud et al., 2007; Ingham et al., 2004; Watkins et al., 2008; Wu et al., 1995). However, given that most BG nuclei are small in size or complex in shape, it is often difficult to determine which nuclei are contributing to activation peaks in the BG (e.g., Watkins et al., 2008). This may explain why, in a recent meta-analysis of functional imaging studies (S. Brown et al., 2005), PWS showed no abnormalities in the BG other than an absence of the weak but reliable

above-baseline observed activation in the left globus pallidus of control subjects during speech.

Despite their limitations in highlighting problems in the BG, imaging studies are very useful in detecting cortical abnormalities. In S. Brown et al.'s (2005) meta-analysis of functional imaging studies, the activation in the left precentral gyrus of persons who do not stutter (PNS) extended into the ventral premotor cortex, while the activation in the left precentral gyrus of PWS was restricted to the ventral primary motor cortex. This result may be interpreted to show that PWS fail to activate the left vPMC, which was indeed confirmed in a recent functional imaging experiment by Watkins et al. (2008). The activation-failures reported by Watkins et al. were observed under conditions in which PWS were dysfluent, suggesting that the degree of deactivation of the vPMC may be causally related to actual moments of stuttering, and not solely to an underlying deficiency. The correlation of vPMC deactivation with actual stuttering is also consistent with observations that the left inferior lateral premotor cortex (ILPrM), which coincides with the vPMC for the large part, is the only left hemisphere motor region in which activation level, for PWS of both sexes, is inversely correlated with stutter rate (Ingham et al., 2004). Also relevant is Neumann et al.'s (2003) finding of abnormally low activation in the left precentral cortex (of which the vPMC is part). In that study, the only dysfluencies produced by the subjects were initial blocks.

There are additional clues supporting the involvement of vPMC in stuttering. The first is that the most common speech elements repeated by PWS are complete syllables or their initial parts. This problem is likely the result of a vPMC malfunction, because this

brain region codes for syllables, as was recently demonstrated using functional imaging techniques (Bohland & Guenther, 2006; Ghosh, Tourville, & Guenther, 2008; Peeva et al., 2009). Another clue is the existence of vPMC-to-vMC projections (Barbas & Pandya, 1987; Kunzle, 1978; Muakkassa & Strick, 1979; Passingham, 1993), which, according to the DIVA model, constitute feedforward or "well-learned" speech motor commands (Guenther et al., 2006). Simulations with the DIVA model showed that impaired readout of feedforward commands, which is one possible outcome of vPMC malfunction, may lead to sound/syllable repetitions (see Chapter 2). Lastly, speech outputs from the BG influence the vPMC via the thalamus (Hoover & Strick, 1993). The vPMC is therefore likely to be affected by the BG problems discussed above. Because the vPMC not only receives projections from, but also sends projections to, the BG (Takada, Tokuno, Nambu, & Inase, 1998), we will refer to this closed circuit by the term *BG-vPMC loop*.

Although suggestive of the BG's involvement in stuttering, the effects of drugs cannot reveal the exact mechanisms of the disorder. Similarly, imaging studies that suggest vPMC involvement in stuttering do not explicate the relation with the apparent BG dysfunction. Most past hypotheses concerning BG dysfunction in stuttering cannot be tested using imaging data, because they cannot predict whether particular brain regions would be over- or under-activated during stuttering (see S. Brown et al., 2005, p. 114). The BG's internal circuits are very complex, including many inhibitory connections, and as a result the relationship between metabolism and signaling is unclear (Alm, 2004, p. 355), making it is difficult to predict neural activations.

Such problems can be partially overcome by testing hypotheses with an extended version of the GODIVA model that includes both cortical and subcortical sites. As a neurobiologically specified model, it is able to predict entire patterns of blood-oxygenation-level-dependent (BOLD) responses observable across the brain regions simulated, much as cerebral and cerebellar activation patterns during fluent speech were predicted based on simulations of its counterpart, the DIVA model (Guenther et al., 2006; Tourville et al., 2008).

This chapter is organized as follows. We first describe the extended GODIVA model, with an emphasis on the newly-developed BG-vPMC loop module. We then use the extended model to simulate neural activity in the brain regions that constitute the BG-vPMC loop, and compare their predicted BOLD responses to published functional imaging data. We also confirm that the model is capable of accounting for the acoustics and kinematics of stuttering, under both normal and medicated conditions. To evaluate the strengths and weaknesses of our hypotheses, each simulation is conducted once with white matter impairment, and once with elevated dopamine levels.

3.2 Neurocomputational modeling of serial speech production

3.2.1 The original GODIVA model

The GODIVA model (Bohland et al., in press) explains how arbitrary utterances which fall within a speaker's language rules can be represented in the brain, and how such utterances can be produced from a finite library of learned motor chunks when activated in the proper order. Based on functional imaging results, previous clinical studies, and previous theoretical models, the GODIVA model simulates the activity in the following

cells: phonological sequence plan- and choice cells that code for sub-syllabic phonological content (phonemes) and reside in the same idealized cortical column of the Inferior Frontal Sulcus (IFS); structural frame representation plan- and choice- cells that code abstract syllable frames and reside in the same column of the presupplementary motor area (pre-SMA); and Speech Sound Maps (SSMs) plan- and choice- cells that code for motor programs and sensory expectations for well-learned syllables (Guenther et al., 2006) and reside in the same column of the vPMC. Each of the cortical regions mentioned is divided into two layers: a *planning* layer (where the plan cells reside), and a *choice* layer (where the choice cells reside). The planning and choice layers correspond to superficial and deep cortical layers, respectively (J. W. Brown, Bullock, & Grossberg, 2004).

Each pair of SSM plan and choice cells corresponds to a specific well-learned syllable. Once the most active SSM choice cell (the cell for the well-learned syllable that comes next) in the vPMC choice layer is selected, it reads out the stored motor programs and sensory expectations appropriate for the next syllable. These instructions are then passed to the DIVA model. To account for stuttering, the functionality of the SSM cells is extended by connecting the vPMC, where these cells reside, with the BG; this forms a BG-vPMC loop. The BG-vPMC loop is separate from the planning loop that connects the BG to the IFS and the pre-SMA. Although both loops pass through the BG, there is no interaction between the two. This chapter is based on simulations of the resulting extended version of the GODIVA model. An illustration of the extended model is shown

in Fig. 3-1. The functionality added to the GODIVA model is described below, followed

by details of its implementation.
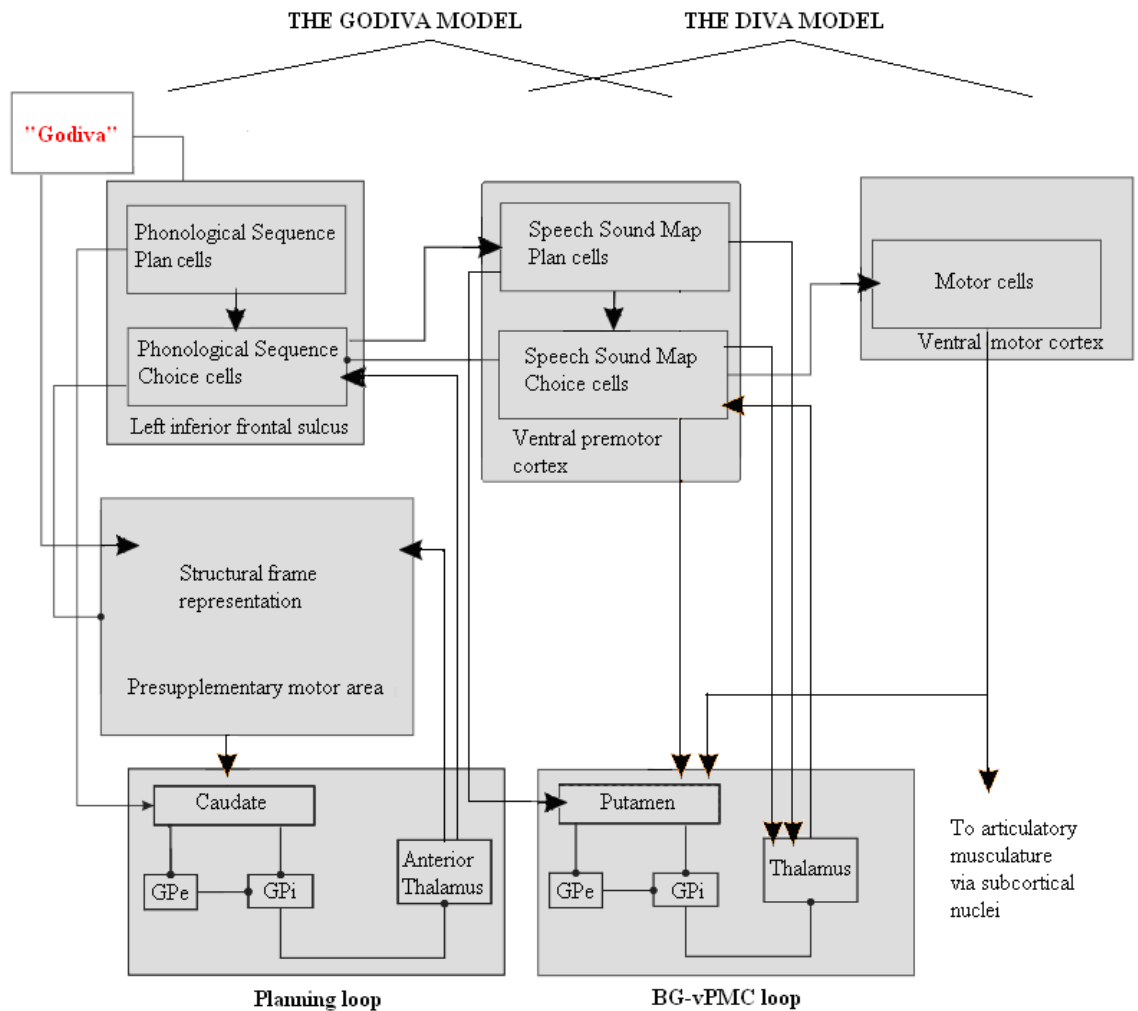
Fig. 3-1. Schematic of the extended GODIVA model and its integration with the DIVA model (only the feedforward control subsystem of the DIVA model is shown). GPi = internal globus pallidus. GPe = external globus pallidus. For simplicity, the details of the presupplementary motor area are omitted.

3.2.2 Extending the GODIVA model to account for stuttering

3.2.2.1 Biasing competition in the cortex

In the GODIVA model, competition via recurrent inhibition among SSM choice cells allows a single sensorimotor program to be chosen for output to the motor apparatus. In general, the information needed to resolve such competitions, and to gate onsets of voluntary actions, is not available in a single cortical patch but is available at BG level due to convergent cortical afferents to zones of the striatum (Bohland et al., in press; J. W. Brown et al., 2004; Passingham, 1993). Therefore, the extended GODIVA model posits a BG loop that originates in the vPMC. Since each BG loop ultimately projects (via the thalamus) to its cortical region of origin, the newly added BG loop is capable of biasing competition in the vPMC choice layer. The BG-vPMC loop accomplishes this by disinhibiting, and thereby exciting, the thalamic cells of the programs most likely to win the competition, i.e., those programs which most closely match the phonological syllable, while at the same time inhibiting the cells of those programs with no chance of winning. This dual function of the BG, namely strengthening desired programs and weakening competing programs, was suggested by Mink (1996, p. 414), and has been explored in numerous computational models. Other than this general effect, the BG-vPMC loop also provides contrast enhancement. The greater the likelihood that a program will win — that is, the more closely it matches the phonological syllable selected in the IFS — the more intensely will it be excited by the BG-vPMC loop.

The BG mechanisms described above provide the cortex with a supplemental form of lateral (or surround) inhibition (see Hikosaka, Takikawa, & Kawagoe, 2000, p.

963), thus allowing the cortex to make more rapid selection once the BG decide that the conditions for releasing a plan are satisfied: increasing lateral inhibition hastens the selection of the SSM choice cell for the correct well-learned syllable. Lateral inhibition provided by the BG, as opposed to that originating in the cortex, gives the central nervous system (CNS) extra degrees of freedom for controlling the extent of lateral inhibition (Gurney, Prescott, & Redgrave, 2001a, 2001b). This BG functionality is strongly modulated by dopamine (Gerfen, 1992; Kitai, Sugimori, & Kocsis, 1976).

The extended GODIVA model posits that within each BG loop there are two pathways, direct and indirect, that fulfill different functions (Mink, 1996, p. 418). The direct pathway has the aforementioned function of biasing competition in the cortex in order to promote the desired program. The indirect pathway's function, on the hand, is to suppress the active motor program based on contextual signals, which permits initiation of the next syllable. This function of the indirect pathway is described next.

3.2.2.2 Initiating the next syllable based on contextual signals

To avoid dysfluencies (e.g., pauses) when switching between syllables, the activity of the SSM choice cell for the current syllable must be quenched predicatively, i.e., before the termination of syllable articulation, in order to allow enough time for the selection of the next syllable. The original GODIVA model employs a non-specific response signal that arrives from vMC and quenches the activity of the currently active SSM cell prior to the

actual completion of articulation (Bohland et al., in press). It is not clear, however, how the vMC is able to predict when syllable articulation is about to terminate[14].

The current version of the GODIVA model provides an alternative method of shifting between syllables, which, as in Alm (2004), assumes BG involvement. Specifically, we assume that the indirect pathway of the BG is responsible for the quenching of the current syllable prior to the termination of articulation. The BG determines when to shift to the next syllable based on information it receives from the cortex. Because the vMC cells, which receive from the vPMC the feedforward commands to be executed (Guenther et al., 2006), project to the striatum (Cowan & Wilson, 1994; Wilson, 2004), the striatum is expected to receive a corollary discharge of ongoing motor commands (cf. M. Parent & Parent, 2006; Wilson, 1995). In speech production, this discharge may be used to anticipate the imminent completion of the current syllable in order to shift to the next syllable in a timely manner (Alm, 2004**,** p. 358). Similar ideas have been entertained for non-speech tasks as well (Deniau, Menetrey, & Charpier, 1996; Ueda & Kimura, 2003).

3.2.2.3   Direct and indirect pathways of the BG-vPMC loop

The idealized circuitry of two channels of the BG, and their corresponding cortical stages, is depicted in Fig. 3-2. To implement the BG-vPMC loop, the GODIVA model hypothesizes projections from thalamic output cells to corresponding SSM choice cells.

---

[14] Bohland et al. (in press) suggest that the inherent delay between sending a motor command and the effect that the motor command has on the articulators permits the motor cortex to predict the termination of articulation. However, it is unlikely that motor cortex by itself can discriminate between early and late phases of the production of a syllable (as opposed to a single articulation).

In addition, following Mink (1996), the GODIVA model hypothesizes that each projection serves to either excite (via supra-baseline thalamic activation) or inhibit (via sub-baseline thalamic activation) the SSM choice cell. The activation of the thalamic cells is determined by a chain of projections (cortex-striatum-pallidum). Although there is significant convergence between these projections in vivo, the model treats them as a set of competitive channels, each represented by two striatal projections neurons (putamen D1 cell of the direct pathway, and putamen D2 cell of the indirect pathway), two striatal interneurons (of the direct and indirect pathways), one internal pallidum cell (GPi) and one external pallidum cell (GPe). The channels interact in the putamen and continue to interact in the putamen's projections to the GPe. In the rest of the BG's nuclei, however, the channels remain independent of one another. Each one of the channels corresponds to a well-learned syllable, and interacts with the idealized cortical column of the vPMC that codes the motor programs and sensory expectations for that syllable.

The direct and indirect pathways originate in the putamen D1 and D2 cells, respectively. Both cell types are modulated by dopamine, and are named after the dopamine receptors they express. Dopamine release strengthens the direct pathway, because the D1 receptor activations on the putamen D1 cells is excitatory to the cell, yet weakens the indirect pathway, because D2 receptor activation by dopamine on the putamen D2 cells is inhibitory to the cell. The direct pathway projects through inhibitory pathways from the putamen to the GPi (Albin, Young, & Penney, 1989), and the indirect pathway projects through inhibitory pathways from the putamen to the GPe (Percheron,

Yelnik, & Francois, 1984), which in turn inhibits the GPi (Mink, 1996; A. Parent & Hazrati, 1995).

For the sake of simplicity, the following discussion will consider each pathway's net effect on the cortex via the thalamus. We propose that because the direct pathway has a net excitatory effect on a single SSM choice cell (J. W. Brown et al., 2004), it is the pathway responsible for biasing cortical competition. The GABA-ergic striatal feedforward interneurons, known to strongly inhibit other striatal cells (J. W. Brown et al., 2004; Koos & Tepper, 1999), are then able to provide the contrast enhancement described above. In contrast to the direct pathway, the indirect pathway has a net inhibitory effect on the thalamus. This permits the indirect pathway to inhibit the SSM choice cell for the current syllable when the configuration of inputs to the striatum allows it to detect that the syllable's articulation is about to terminate. Because the indirect pathway acts in a diffuse manner, concurrently inhibiting all thalamic cells (Mink, 1996), the inhibition is not specific to the current syllable. However, since the SSM choice cell for the current syllable is the only one active during syllable articulation, non-specific inhibition is appropriate. The model's supposition that the indirect pathway is differentially excited by a corollary discharge from vMC cells fits with Lei et al.'s (2004) findings that collaterals of deep-layer motor cortex's efferents synapse preferentially on the striatal cells of the indirect pathway. These collaterals thus are well-suited to provide the indirect pathway with corollary discharge of motor commands (cf. Cowan & Wilson, 1994).
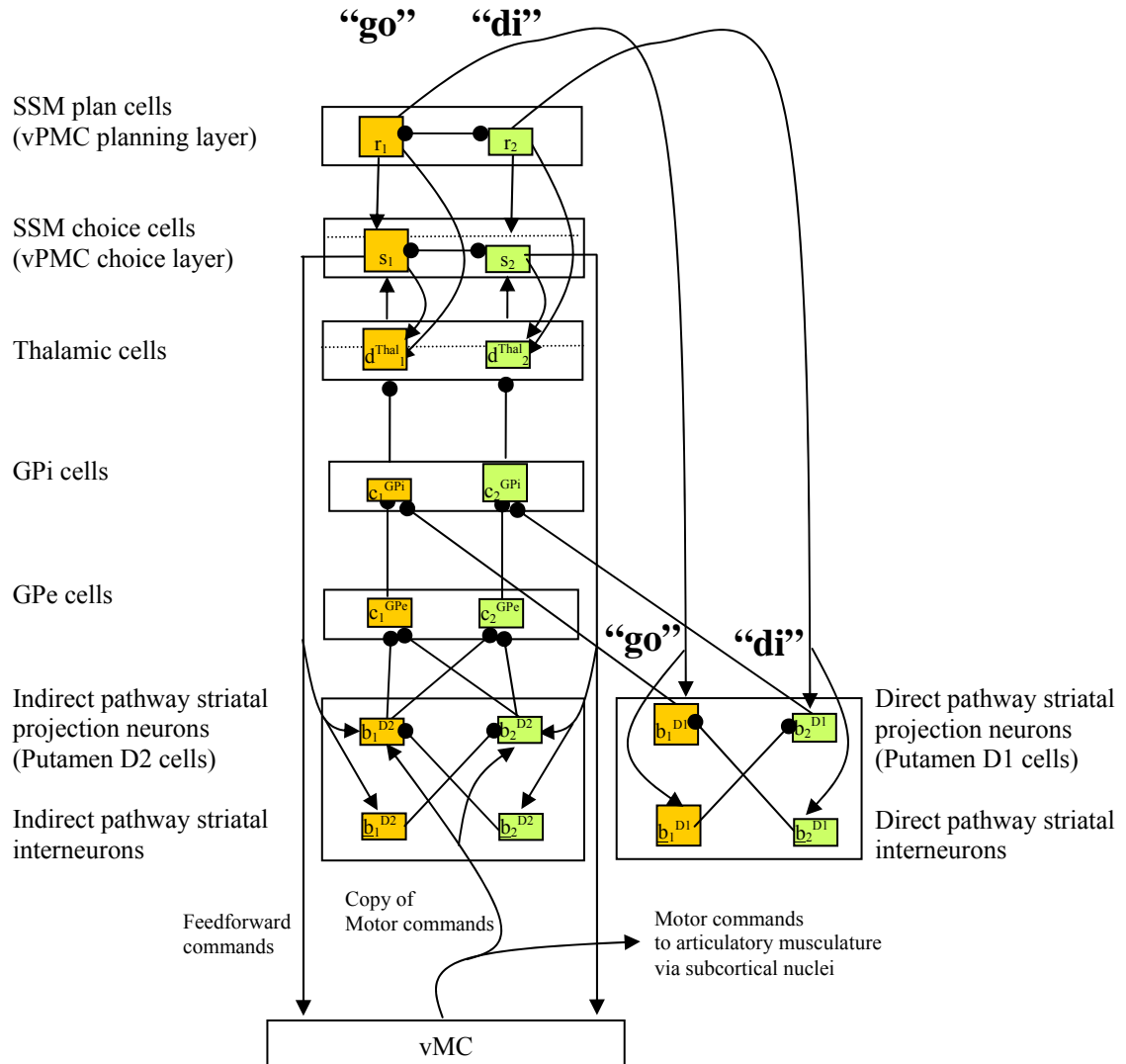
Fig. 3-2. Idealized circuitry of the BG-vPMC loop combined with typical activation levels of the circuit's cells. The model treats the cortico-striatal-pallidal projections as a set of competitive channels. Two of these are included in the diagram: the first channel for the well-learned syllable "go" (yellow), and the second channel for the well-learned syllable "di" (green). Within each channel are: two striatal projection neurons (putamen D1 cells of the direct pathway shown projecting to GPi, and putamen D2 cells of the indirect pathways shown projecting to GPe), two striatal interneurons (direct and indirect) shown as mediators of lateral inhibition, one internal pallidum cell (GPi), and one external pallidum cell (GPe). The cortical columns are shown as well, each represented by one SSM plan cell (r_i at the vPMC planning layer) and one SSM choice cell (s_i at the vPMC choice layer). The cells are depicted as bars, with bar heights representing each cell's level of activation when the GODIVA model is about to produce the syllable "go" in the sequence "go.di. və".

To produce a syllable, the direct and indirect pathways interact as follows. First, the direct pathway biases competition in the vPMC in favor of the correct syllable. As soon as a SSM choice cell is selected in the vPMC, the cell starts reading out the feedforward commands into the vMC cells targeted by its efferents. The selected SSM choice cell also primes the specific indirect pathway striatal projection neuron whose role is to detect the completion of the selected syllable. Next, the vMC starts executing the motor commands; a process that continues until the primed striatal neuron, which constantly receives copies of executed commands, detects a combination of commands that indicates that the syllable is about to terminate. The primed striatal neuron will then activate the indirect pathway that immediately comes into play. By terminating the current syllable, the pathway's activation clears the way to the production of the next syllable in line.

Although the strongest activation in the model's indirect pathway occurs before a syllable shift, the indirect pathway is moderately active during syllable selection as well (before the first syllable in the sequence or between syllables). Because it is targeted by the vPMC choice layer, the indirect pathway is able to detect the number of SSM choice cells in competition there, and then exerts appropriate level of inhibition on the thalamus. This inhibition is crucially important in ensuring that the thalamus properly expresses the gradient in the activation levels of the SSM choice cells, i.e., the contrasts among the well-learned syllables competing for execution. When the contrast in the thalamus is lost, the thalamus excites all SSM choice cells to the same degree, and selection of the motor program for the next syllable cannot be expedited.

3.2.2.4   Interfacing the extended GODIVA model with the DIVA model

The GODIVA model simulates high-level processes in the speech production system, but does not address the speech output stage. For the GODIVA model to generate acoustic and kinematic output, it had to be integrated with the DIVA model (as suggested by Bohland, 2007). The DIVA model is capable of generating acoustic and kinematic outputs for fluent speech (Guenther, 1995, 2006; Guenther et al., 2006; Guenther et al., 1998; Nieto-Castanon et al., 2005; Tourville et al., 2008), and, as was recently demonstrated, for dysfluent speech as well (Chapter 2). For the purposes of this study, the DIVA model was used as-is (Fig. 3-1), with one minor modification: we assume that the SSM choice cell activation level required to properly read out feedforward commands for a given syllable is higher than the activation level required to read out the corresponding sensory expectations.

While most information flow in the combined model is from the extended GODIVA model to the DIVA model (i.e., flowing from brain regions that handle high-level speech processing to brain regions that handle low-level processing), there is one case in which the DIVA model returns information to the GODIVA model: a "reset" of current-syllable production when faced with an error too big to be corrected via feedback control (not shown in Fig. 3-1). This "reset" is implemented as a quenching of the activity in the vPMC choice layer, starting shortly after error detection and ending when the next attempt to produce the syllable is initiated.

3.2.3 Extended GODIVA model specifications

3.2.3.1 Implementation

The extended GODIVA model, like the original GODIVA model, is implemented computationally as a neural network governed by a set of ordinary differential equations that capture key features of membrane dynamics, including the bounded range of cell potentials (e.g., Grossberg, 1973, 1978). This approach permits computation of transient as well as steady-state network behavior, and allows model simulations to be constrained by neural activity data collected from monkeys performing non-speech motor-sequencing tasks. The extended GODIVA model uses all of the equations of the original GODIVA model as-is (as specified in Bohland et al., in press), except equation (14). Equations (1)-(13), which are shared between the models, are given in appendix A. The main variables of the equations are given in appendix B. Table 3-1 lists the variables used to specify the BG-vPMC loop in the extended GODIVA model.

Table 3-1. Legend of symbols used to refer to cell populations in the specification of the BG-vPMC loop of the extended GODIVA model.

Cell Group	Symbol
SSM Plan Cells (vPMC planning layer)	r
SSM Choice Cells (vPMC choice layer)	s
BG-vPMC Loop Direct Pathway Striatal Projection Neurons (Putman D1 cells)	b^{D1}
BG-vPMC Loop Direct Pathway Striatal Interneurons	$\underline{b}^{D1}$
BG-vPMC Loop Indirect Pathway Striatal Projection Neurons (Putman D2 cells)	b^{D2}
BG-vPMC Loop Indirect Pathway Striatal Interneurons	$\underline{b}^{D2}$
BG-vPMC Loop GPi Cells	c^{GPi}
BG-vPMC Loop GPe Cells	c^{GPe}
BG-vPMC Loop Thalamic Cells	d^{Thal}

3.2.3.2 Ventral premotor cortex

Each idealized column of the vPMC represents one well-learned syllable. The activity in the plan cells indicates the degree of match between the set of active phonological cells in the IFS choice layer (the forthcoming phonological syllable) and the stored sensorimotor programs associated with the SSM columns. The match is computed via an inner product of the IFS choice-layer inputs with synaptic weights, which are "hard-wired" such that correct well-learned syllable programs will have the highest activation in the planning layer of the vPMC. Nevertheless, other syllable programs which have a partial match to the phonological syllable will be active as well, although to a lesser degree. The activity of the SSM plan cell, r_k, is further described in Bohland et al. (in press), and is given in equation (13).

The SSM plan cell r_k gives specific excitatory input to the SSM choice cell s_k within the same idealized cortical column. The activation of s_k in the current model is given by:

$$\dot{s}_k = A_s s_k + \left(B_s - s_k\right) y\left(d_k^{Thal}, s_k\right) I - s_k \left(\sum_{j \neq k} y\left(d_j^{Thal}, s_j\right) + \Omega_{reset} \right) \qquad (14)$$

The dynamics are such that the most active well-learned syllable's program in the planning layer r_k is selected in the choice layer (Bohland et al., in press). This is due to winner-take-all competition in the choice layer: only one cell "wins" the competition, while all the "loser" cells are quenched. This competition is mediated by self-excitation (the term $y\left(d_k^{Thal}, s_k\right)$) and surround inhibition (the term $y\left(d_j^{Thal}, s_j\right)$). Self-excitation also ensures that the "winner cell" will maintain its activation regardless of any changes

in its inputs. *I* is a binary [0 or 1] input arriving from brain regions not modeled by the extended GODIVA model; when set to 1, it initiates the production of the first syllable in the sequence. If necessary, resetting it to 0 will stop the production of the sequence before it is over. Ω_{reset} is a suppression signal which arrives from the DIVA model when a "reset" is initiated (as described in Section 3.2.2.4).

Both the self-excitation and the surround-inhibition terms in (14) use a faster-than-linear function modulated by input from the thalamus. It is defined by:

$$y(d^{Thal}, s) = z\left(\left[d^{Thal} - T_d\right]^+\right) s^{1.5} \qquad (15)$$

where the thalamic input d^{Thal} is thresholded by T_d and strongly amplified via a faster-than-linear activation function, $z(x) = x^4$ (cf. Grossberg, 1973), whose neural correlate may be a spike rate that varies non-linearly with membrane potential. The modulation by the thalamus biases competition in the vPMC choice layer in favor of likely winners, as discussed in Section 3.2.2.1. Cortical cells whose corresponding thalamic cells' activation levels are below T_d are totally deprived of self-excitation, and will quickly be eliminated from competition. Cortical cells whose thalamic cell activation levels are above T_d will be excited in proportion to the firing rate of the thalamic cell.

When the competition yields a "winner" SSM cell in the choice layer, the cell excites its corresponding thalamic cell, which then maintains the SSM choice cell's self-excitation. In the absence of a reset signal Ω_{reset}, the deletion of this SSM choice cell activity is possible only when the thalamic cell is pushed below the threshold T_d by inhibition from the indirect pathway. Because the natural decay of SSM choice cells is

not strong, it might be that stopping self-excitation is not sufficient to quickly terminate choice cell activation. However, the baseline lateral inhibition in the choice layer ensures prompt termination of activity.

Finally, binary activation of SSM choice cells in the most recent instantiation of the DIVA model (Guenther et al., 2006) makes it difficult to integrate it with the GODIVA model and to simulate impaired readout of feedforward commands. Here we replaced DIVA's SSM choice cells, $P_k(t)$, with a function that reflects GODIVA's SSM choice cell activation levels:

$$P_k(t) = \begin{cases} 1 & \text{if } s_k(t) > T_{Feedforward} \\[2em] \dfrac{s_k(t) - T_{Feedback}}{T_{Feedforward} - T_{Feedback}} & \text{if } T_{Feedforward} > s_k(t) > T_{Feedback} \qquad (16) \\[2em] 0 & \text{otherwise} \end{cases}$$

where $s_k(t)$ is the GODIVA's SSM choice cell, $T_{Feedforward}$ is a threshold that ensures that feedforward commands are read out in full only after the SSM choice cell wins the competition in the choice layer of the vPMC, and $T_{Feedback}$ is a threshold above which only weak feedforward commands are read out. Readout of feedforward commands, whether weak or strong, is always accompanied by readout of their sensory expectations.

3.2.3.3 Putamen

The activity of the direct pathway striatal projection neurons (putamen D1 cells) in BG channel j is given by:

$$\dot{b}_j^{D1} = -A_{D1}b_j^{D1} + \left(\beta_{D1}B_{D1} - b_j^{D1}\right)r_j - b_j^{D1}\left(\sum_{k \neq j}\underline{b}_k^{D1}\right) \qquad (17)$$

where r_j is the SSM plan cell activation that provides input to the BG loop, and the product $B_{D1}\beta_{D1}$ controls the maximum strength of the excitatory term. Whereas B_{D1} is a constant, β_{D1} is determined by the amount of dopamine that binds to the D1 receptors of the putamen (cf. Frank, 2005), which has the value of 1 when dopamine levels are normal. Because SSM plan cells are the main inputs to the putamen D1 cells, the activations of the later cells reflect the pattern of activity in the former ones. This pattern of activity depends on the next syllable in the sequence: The SSM plan cells, and thus the putamen D1 cells, that are active at each phase of the sequence represent the set of similar syllables from which the next syllable is selected. This property of the model is consistent with the recent findings that cells in the putamen are selective to the phase of the sequence (Filatova, Orlov, Tolkunov, & Afanas'ev, 2005; Ueda & Kimura, 2003).

The cell b_j^{D1} also receives feedforward inhibition from striatal interneurons $\underline{b}_j^{D1}$ in the other BG channels ($k \neq j$). The activity of a direct pathway striatal inhibitory interneuron in channel j is governed by:

$$\dot{\underline{b}}_j^{D1} = -A_{D1}\underline{b}_j^{D1} + \left(B_{D1} - \underline{b}_j^{D1}\right)r_j \qquad (18)$$

As in the case of the projection neuron, the input r_j is a SSM plan-cell activity. In contrast with the projection neurons, however, model interneurons are not inhibited by other striatal cells (see J. W. Brown et al., 2004)

The activity of the indirect pathway striatal projection neuron (putamen D2 cell) in BG channel j is given by:

$$\dot{b}_j^{D2} = -A_{D2}b_j^{D2}$$
$$+\left((1/\beta_{D2})B_{D2}-b_j^{D2}\right)\left(\lambda_{WMF}m_j\left(\dot{M}\right)f\left(\left[s_j-T_{Feedforward}\right]^+\right)+f\left(\left[s_j-\theta_s\right]^+\right)\right) \qquad (19)$$
$$-b_j^{D2}\left(\sum_{k\neq j}b_k^{D2}\right)$$

where β_{D2} and B_{D2} control the maximum strength of the excitatory term. B_{D2} is constant, whereas β_{D2} represents the amount of dopamine that binds to the D2 receptors of the putamen. As with β_{D1}, β_{D2} equals 1 at normal dopamine levels. In contrast to the direct pathway, increased binding in the indirect pathway weakens the excitatory term (cf. Frank, 2005). θ_s is a low threshold on SSM choice cell activity, s_j, and $f(x)$ is a binary function defined by:

$$f(x) = \begin{cases} 1 & if\ x > 0 \\ 0 & otherwise \end{cases} \qquad (20)$$

Term m_j in (19) is a function that detects when articulation of the syllable coded by SSM choice cell j is about to terminate. In particular, it calculates the match between the motor commands $\dot{M}$ currently executed by vMC (see Guenther et al., 2006), as a result of activating SSM choice cell $j,$ and a long term memory for the motor command state that predicts imminent termination of this specific syllable. If the match is high, such that the articulation of the syllable is about to terminate, then $m_j(\dot{M}) = 1$, and the putamen D2 cell becomes strongly active. This strong activation ultimately inhibits thalamus and that undercuts SSM cell activation. While the match remains low, i.e.,

$m_j(\dot{M}) = 0$, the syllable is still far from terminating, and putamen D2 cell activation

remains low. Because compact regions in the putamen receive corollary discharges of

motor commands via collaterals of vMC's efferents, the strength of the striatal activation

generated during a match depends on the corticostriatal projection's integrity, which is

quantified by λ_{WMF} in the above formula.

The model does not attempt to explicate learning of long term memories for states

that predict imminent terminations. Instead, m_j is algorithmically calculated using the

following formula:

$$m_j(\dot{M}) = \max\left(\left((H_j^M - L_j^M) - abs(\dot{M} - L_j^M) - abs(\dot{M} - H_j^M) + 1\right), 0\right) \qquad (21)$$

where L_j^M and H_j^M code for the low and high bounds, respectively, of a sixteen

dimensional region associated with the syllable coded by the idealized vPMC column j.

Each motor command, $\dot{M}$ (as defined in section 2.2), can be described by sixteen values,

corresponding to the commands sent to the various articulators. Any motor command that

falls inside this region signals this syllable imminent termination. Matching the current

motor command with a region, instead of a single point in the motor commands space,

keeps the model robust.

To prevent putamen D2 cells of other syllables from mistakenly detecting

completion based on the motor commands of the current syllable, the function m_j in (19)

is multiplied by the term $f\left(\left[s_j - T_{Feedforward}\right]^+\right)$ that ensures that only the current

syllable's putamen D2 cell can become strongly active. The SSM choice cell of the

current syllable, $s_{current}$, is the only one with activation above $T_{Feedforward}$, thus only

$$f\left(\left[s_{j=current} - T_{Feedforward}\right]^{+}\right)$$ is non-zero.

Lastly, the term $f\left(\left[s_j - \theta_s\right]^{+}\right)$ in (19) enables weak activation of the indirect

pathway during syllable selection. It represents excitation of indirect pathway striatal

projection neurons by all SSM choice cells with supra-threshold activity. This scales the

inhibition that the indirect pathway exerts on the thalamus (see Section 3.2.2.3), such that

the total inhibition is proportional to the number of SSM cells competing in the vPMC

choice layer.

Like the putamen D1 cells, also the putamen D2 cells, b_j^{D2}, receive feedforward

inhibition from striatal interneurons $\underline{b}_j^{D2}$ in the other BG channels ($k \neq j$). The activity

of an indirect pathway striatal inhibitory interneuron in channel j is governed by:

$$\dot{\underline{b}}_j^{D2} = -A_{D2}\underline{b}_j^{D2} + \left(B_{D2} - \underline{b}_j^{D2}\right)f\left(\left[s_j - T_{Feedforward}\right]^{+}\right) \qquad (22)$$

$T_{Feedforward}$ is equal to the vPMC choice layer selection threshold, i.e., a SSM cell needs to

exceed this threshold in order to fully read out the feedforward commands (as defined in

Section 3.2.3.2). When a SSM choice cell exceeds this threshold, it will also inhibit, via

the indirect-pathway interneuron in the corresponding BG channel, the putamen D2 cells

in the other channels. This ensures that the indirect pathway striatal projection neurons

that assisted in inhibiting the thalamus during the selection period (the minor role of the

indirect pathway in the extended GODIVA model) will not interfere with the detection of

syllable termination (the major role of the indirect pathway in the model).

3.2.3.4 Globus pallidus

The activation of pallidal cells, like that of the striatal cells, is dependent on the phase of the sequence (Mushiake & Strick, 1995). The putamen D1 cells are those informed by the cortex regarding the next element in the sequence; information which they then transmit to the GPi. The putamen D1 projection cells connect to GPi cells within the same BG channel via an inhibitory synapse. The activity of the GPi cell c_j, which is itself inhibitory to a corresponding thalamic cell d_j^{Thal}, is given by:

$$\dot{c}_j^{GPi} = \beta_{GPi}\left(B_{GPi} - c_j^{GPi}\right) - c_j^{GPi}\left(b_j^{D1} + 1.46\left(c_j^{GPe}\right)\right) \qquad (23)$$

where β_{GPi} and B_{GPi} control the level of spontaneous tonic activation of the GPi cell. The tonic activation of the GPi (Delong, 1971) is essential to the proper functioning of the model. It permits changes of GPi cell activation to mediate both net excitation and inhibition of the thalamic cell. When the putamen is silent, the tonic activation of the GPi cell moderately inhibits the thalamus. When the GPi cell is inhibited by the putamen D1 cell, its activation drops below tonic levels, and the thalamic cell it targets is excited due to disinhibition. When the tonic inhibition of the GPi cell by the GPe is removed, the GPi cell activation rises above tonic levels, and the thalamic cell it targets is inhibited.

Putamen D2 cells connect to GPe cells within the same BG channel via an inhibitory synapse. The activity of the GPe cell c_j^{GPe}, which is itself inhibitory to the GPi cell within the same channel, is given by:

$$\dot{c}_j^{GPe} = \beta_{GPe}\left(B_{GPe} - c_j^{GPe}\right) - c_j^{GPe}\left(10\sum_k b_k^{D2}\right) \qquad (24)$$

where β_{GPe} and B_{GPe} control the level of spontaneous tonic activation of the GPe cell, and the inhibitory inputs in the last term arrive from putamen D2 cells. Activity in each one of the putamen D2 cells b_k^{D2} is modeled to inhibit all GPe cells because the projections from the putamen to GPe are non-specific (Flaherty & Graybiel, 1994; Percheron et al., 1984).

At rest the GPe is tonically active, moderately inhibiting the GPi. However, when one of the putamen D2 cells is activated, the GPe is inhibited, and the degree of inhibition it exerts on the GPi will be reduced. Thus, the putamen D2 cells can excite (by disihibition) the GPi via the GPe. This excitation, however, is transient: new feedforward commands are constantly being sent to the articulators, and they fall within the "match" region for only a short time. Such transient GPi excitation can explain the spike in pallidal activity which precedes all but the first in a sequence of hand movements (Brotchie, Iansek, & Horne, 1991). Because the GPi inhibits the thalamus, this spike will be translated to transient inhibition of the thalamus.

To summarize, in the extended GODIVA model the BG are capable of both exciting (disinhibiting) and inhibiting the thalamus (Mink & Thach, 1993; Wichmann & Delong, 1996). The putamen D1 cells excite thalamic cells via the GPi (the direct pathway), while the D2 putamen cells inhibit thalamic cells via the GPe-GPi route (the indirect pathway).

3.2.3.5 Thalamus

The activity of a thalamic cell d_k^{Thal} is given by:

$$\dot{d}_k^{Thal} = \left(B_d - d_k\right)\left(r_k + \left[d_k^{Thal} - \theta_d\right]^+ + z\left(s_k\right)\right) - 10 d_k^{Thal}\left(c_k\right) \qquad (25)$$

where r_k is the activation of the SSM plan cell in idealized column k, s_k is the activation of the SSM choice cell in idealized column k, and c_k is the activation of the GPi cell in the same BG channel. Because each thalamic cell excites itself (d_k^{Thal} is included in the excitatory term), all thalamic cells initially excited by the cortex will ultimately converge, to a constant activation level. During speech, however, the putamen is not silent. The putamen cells are excited by the cortex, and modulate thalamic cells by using inhibition and disinhibition via the GPi. The last excitatory input to the thalamus, $z(s_k)$, arrives from the SSM choice cells in the vPMC. Being much faster-than-linear, it ensures that activity in the thalamic cells of winning cortical cells persists until inhibited by the indirect pathway. This inhibition must push the thalamic cell activation below θ_d in order to ensure its deactivation. Our proposal that projections to the thalamus from superficial (planning) and deep (choice) layers differ in nature is based on recent evidence from Zikopoulos & Barbas (2007).

3.2.3.6 Prediction of BOLD responses

As stated in the introduction, one goal of the current modeling work is to generate predictions of entire patterns of blood-oxygenation-level-dependent (BOLD) responses associated with stuttering in different speaking conditions. To this end, we have identified likely neuroanatomical locations of the model's components (Guenther et al., 2006), which allow us to run "simulated functional Magnetic Resonance Imaging (fMRI) experiments" where the model produces speech sounds in both medicated and

unmedicated conditions. Based on the model cell activities, we then generated simulated hemodynamic response patterns.

The simulated hemodynamic responses were produced along the lines of the procedure described in Guenther et al. (2006). First of all, model cell activities were thresholded; only activities higher than the threshold are assumed to generate hemodynamic response. Based on an additional assumption that cell activity contributes exponentially to hemodynamic response, the resultant activity was squared. The squared activity of each cell was then convolved with an idealized hemodynamic response function that approximates the transformation from cell activity to hemodynamic response in the brain (for details, see Guenther, 2006). Lastly, the overall hemodynamic response for each brain region was calculated by summing the responses of the region's cells. R_{BOLD}^{x}, the hemodynamic response of the brain region that includes the x_i cells (where x can be substituted for any of the cell types in the model) is thus given by:

$$R_{BOLD}^{x} = \sum_{i} f_{HRF}\left(\left[f_{Habituate}\left(x_i - T_{BOLD}^{x}\right)\right]^2\right) \qquad (26)$$

where f_{HRF} is the idealized hemodynamic response function, and T_{BOLD}^{x} is the threshold for generating BOLD responses for cells of type x. For the thalamic cells, this threshold is assumed to be high, thus set to half of the cells' dynamic range ($T_{BOLD}^{d^{Thal}} = 5$), while for the other cells it is assumed to be low, thus set to 0. The function $f_{Habituate}$ permits a more realistic prediction of BOLD responses. It accounts for the assumption that cells that fire at their maximum rate for extended periods of time are assumed to "habituate" in their ability to recruit additional blood flow. As a result their contribution to the BOLD

response tapers off with time. In the current implementation, this "habituation" begins once a cell has maintained maximum-level firing rate for 200ms. However, there is no corresponding habituation in the cell's input or responses to inputs.

3.3 Results

Computer simulations were performed to test our two hypotheses about the etiology of stuttering. One involved elevated dopamine levels (affecting both the D1 and D2 receptors of the putamen), and the other involved impaired white matter projection (in the corticostriatal fibers that arise from collaterals of vMC efferents). To access each hypothesis, we modified one of the critical parameters of the model (listed in appendix C) while keeping the other parameter values constant. For each parameterization, we then report the time course of activity in several key model components, as well as the BOLD response predicted for that activity. We run the simulations once when the extended GODIVA model produces its name (/godivə/) at normal conditions, and once when it

produces its name under the influence of D2 antagonists. The behavioral outcome of the simulations depends on the selection of the $T_{Feedforward}$ and $T_{Feedback}$ thresholds. The simulations reported here focus on speaking blocks. The model's further applicability to repetitions and prolongations is treated in the discussion.

In order to make predictions about differences in brain activations between PWS and PNS, we also simulated the normal (unimpaired) version of the extended GODIVA model. All simulations assume that the sensorimotor programs for the 300 most frequent syllables from the CELEX database (which include all of the syllables in the target

words) were acquired by the DIVA portion of the circuit and are represented by SSM cells presumed to reside in the vPMC.

3.3.1   Fluent speech

The activity in the BG-vPMC loop during fluent production of the sequence "go.di.və" by an intact version of the GODIVA model (*normal GODIVA version*) is shown in Fig. 3-3. The "input" to this simulation is a gradient set of parallel pulses from higher-order linguistic areas. These inputs create an activity gradient across the /g/, /d/, and /v/ phoneme cells in syllable position 3 (onset consonants) and a gradient across the /o/, /i/, and /ə/ phoneme cells in syllable position 4 (vowel nucleus) in the IFS planning layer, as well as a gradient across three "copies" of the [CV] frame cell in the pre-SMA (not shown). The processing of these inputs results in co-temporal activation of IFS choice cells coding /g/ and /o/, which then drive activity in the model's vPMC planning layer which contains units corresponding to 300 well-learned syllables. Multiple cells representing sensorimotor programs for syllables become active, each partially matching the phonological sequence representation in the IFS choice layer. The color of each cell activation in Fig. 3-3 indicates the syllable to which the cell corresponds. The most active of these SSM plan cells (whose activation is plotted in blue) codes for the best matching syllable (in this case, "go"). The tonic input that enables selection of SSMs in the vPMC is applied from t=540ms and on. Activity of the SSM choice cells is then observed, modulated by input from the thalamus, to be described next.
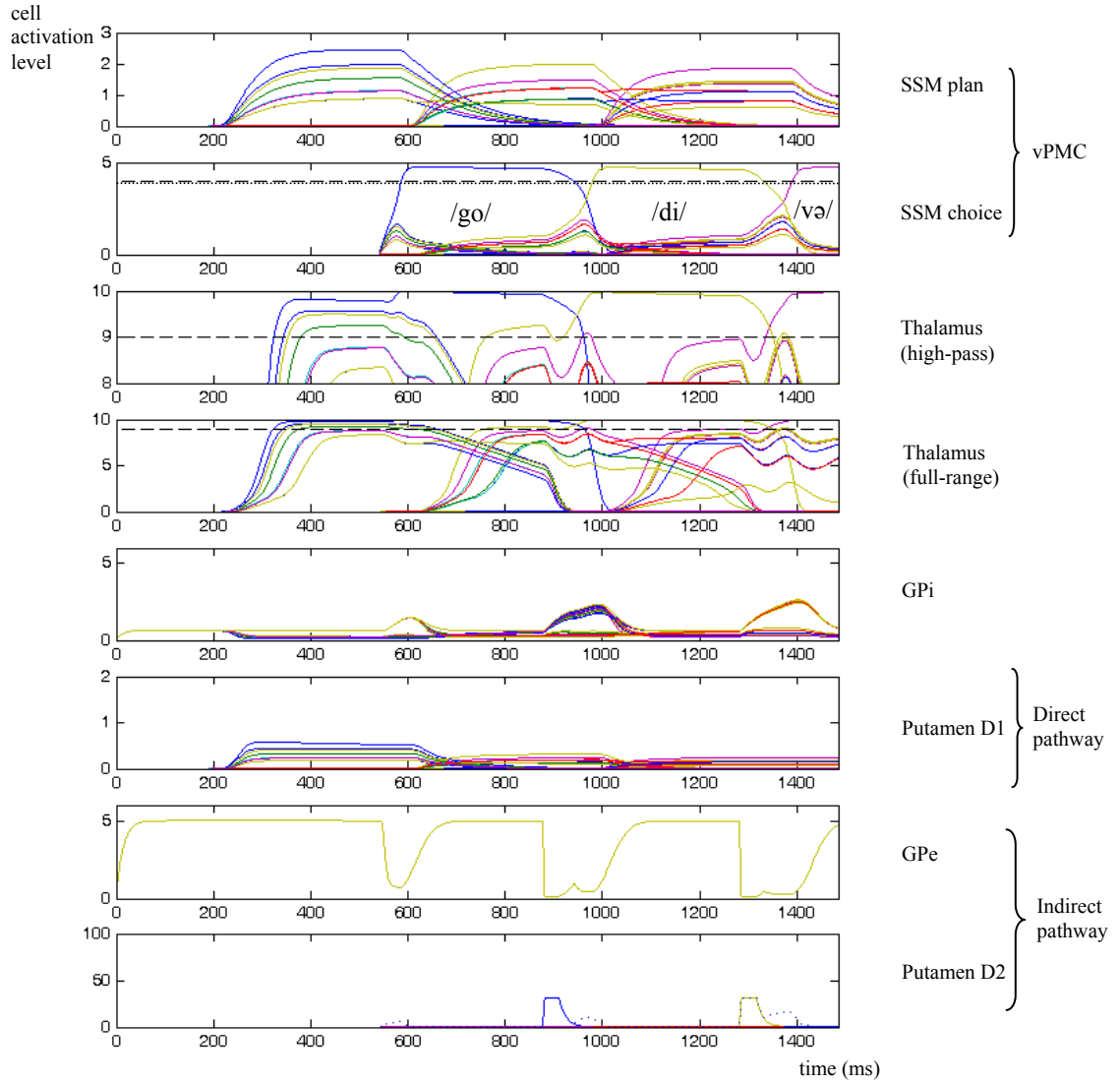
Fig. 3-3. Time course of neural activity in the BG-vPMC loop. Each idealized cortical column represents a specific well-learned syllable, with activations of the corresponding SSM plan cell and SSM choice cell plotted in the same color. Also shown in the same color are the activations of the cells in the BG channel that correspond to the same syllable: the thalamic cell, the GPi cell, the direct pathway striatal projection neuron (putamen D1 cell), the GPe cell, and the indirect pathway striatal projection neuron (putamen D2 cell). The dots in the bottom plot indicate the activity of all putamen D2 cells combined. Dashed straight lines indicate the $T_{Feedforward}$ threshold in the vPMC choice layer, and the T_d threshold in the thalamus. Dotted straight line indicates the vPMC choice-layer $T_{Feedback}$ threshold, which is just below the $T_{Feedforward}$ threshold. For clarity, the thalamic cell activations are shown both in high-pass and full-range views.

The putamen D1 cells, driven by the input from the SSM plan cells, reflect the pattern of activity in the vPMC planning layer. Soon enough, the putamen D1 cells inhibit the GPi cells, driving the later below their tonic level of activation. The weakening of the GPi cells (which is proportional to the strength of activation in their corresponding putamen D1 cells) removes the inhibition they usually exert on the thalamus, which now, like the putamen, reflects the pattern of activity in cortex (with the cell for "go" having the highest activation).

As can be observed in the figure, all well-learned syllables that bid for execution have their thalamic cells active to some degree. However, only those thalamic cells whose corresponding SSM plan cells have the strongest activations in the vPMC exceed the thalamic threshold. These thalamic cells enable self-excitation in their target SSM choice cells, thus allowing the SSM choice cells to participate in the syllable-selection process. Because thalamic cells modulate cortical self-excitation in proportion to their level of activation, the SSM choice cell for "go", whose thalamic cell has the highest activation in the thalamus, has the strongest self-excitation. This helps the SSM choice cell for "go" to quickly win the cortical competition. Once it exceeds the threshold $T_{Feedforward}$, the cell reads out the motor commands for the syllable "go", and due to its high activation, also returns strong input to its corresponding thalamic cell. This is the reason that the thalamic cell for "go" persists at maximum activation, while the other thalamic cells decay.

At t=877ms, a phasic activation can be observed in the putamen D2 cells of the indirect pathway. This results in a deep pause in the firing of all GPe cells, and

consequently a phasic increase in the activity of all GPi cells. This chain ultimately leads to a phasic inhibition that quenches all the thalamic cells, including the cell for the active SSM choice cell for "go". The only thalamic cells spared are those that are excited via the direct pathway — in this case, the thalamic cells corresponding to the syllables which best match the phonological syllable /di/. At this point, another syllable-selection process takes place and the SSM choice cell for "di" is selected. The last syllable in the sequence, "və", is likewise selected and terminated in the manner described above.

The dip in GPe activity and peak in GPi activity around t=573ms are attributable to the weak activation of many putamen D2 cells at approximately the same time. The putamen D2 cells contribute to the modulation of the thalamus during the syllable selection period, as explained in section 3.2.3.3. Because the roles of the GPe and GPi cells in the model, both when selecting and when deleting SSM choice cell, are limited to relaying and integrating information between the putamen and the thalamus, their activations will not be reported for the rest of the simulations.

3.3.2   Dysfluent speech due to elevated dopamine levels

Fig. 3-4 shows the time course of activity in several key model components in simulations of a version of the extended GODIVA model with elevated dopamine levels (*DA GODIVA version*). In this version, both dopamine binding parameters β_{D1} and β_{D2} are set by default to 1.75. Panel (a) shows a simulation of the model producing dysfluent speech. A comparison to the performance of the normal GODIVA version (Fig. 3-3) reveals that the unmedicated DA GODIVA version reads out the sensorimotor program

for the syllable "go" too late. The upper blue line, indicating the activity of the SSM choice cell for "go", exceeds threshold at t=810ms, compared to t=587ms for the normal GODIVA version. Based on the parameter choice for this simulation, the behavioral correlate of the delayed SSM choice cell activation is a block (lasting for 317ms) on the first syllable of the word /godivə/.

Another notable deviation in the simulation of the unmedicated DA GODIVA version is in the increased activities of all putamen D1 cells, which is to be expected with higher dopamine levels. While not affecting the ordering of cell activations in the putamen, the increased activation in the direct pathway does affect the order of cell activations in the thalamus: throughout the selection period for the first syllable in the sequence, there are at least 2 thalamic cells with equal maximum activation level. One of these cells is the thalamic cell that corresponds to the well-learned syllable "go". Because it does not have an advantage over its runner-up, it cannot bias cortical competition in favor of the SSM choice cell for "go".

Fig. 3-4(b) shows a simulation in which the DA GODIVA version is mostly fluent when under the influence of D2 antagonists. Treatment with D2 antagonists was simulated as a tenfold decrease in binding of dopamine to D2 receptors, but with no change in binding of dopamine to D1 receptors (β_{D1}=1.75, β_{D2} = 0.175). The simulation shows that much like the normal GODIVA version, the readout of the sensorimotor program for the well-learned syllable "go" starts relatively quickly (t=615ms). The behavioral correlate is a short block (28ms long) on the first syllable of the word /godivə/.

(a) Unmedicated DA GODIVA: Cell activities

(b) Medicated DA GODIVA: cell activities

(c) Unmedicated DA GODIVA: BOLD responses

(d) Medicated DA GODIVA: BOLD responses
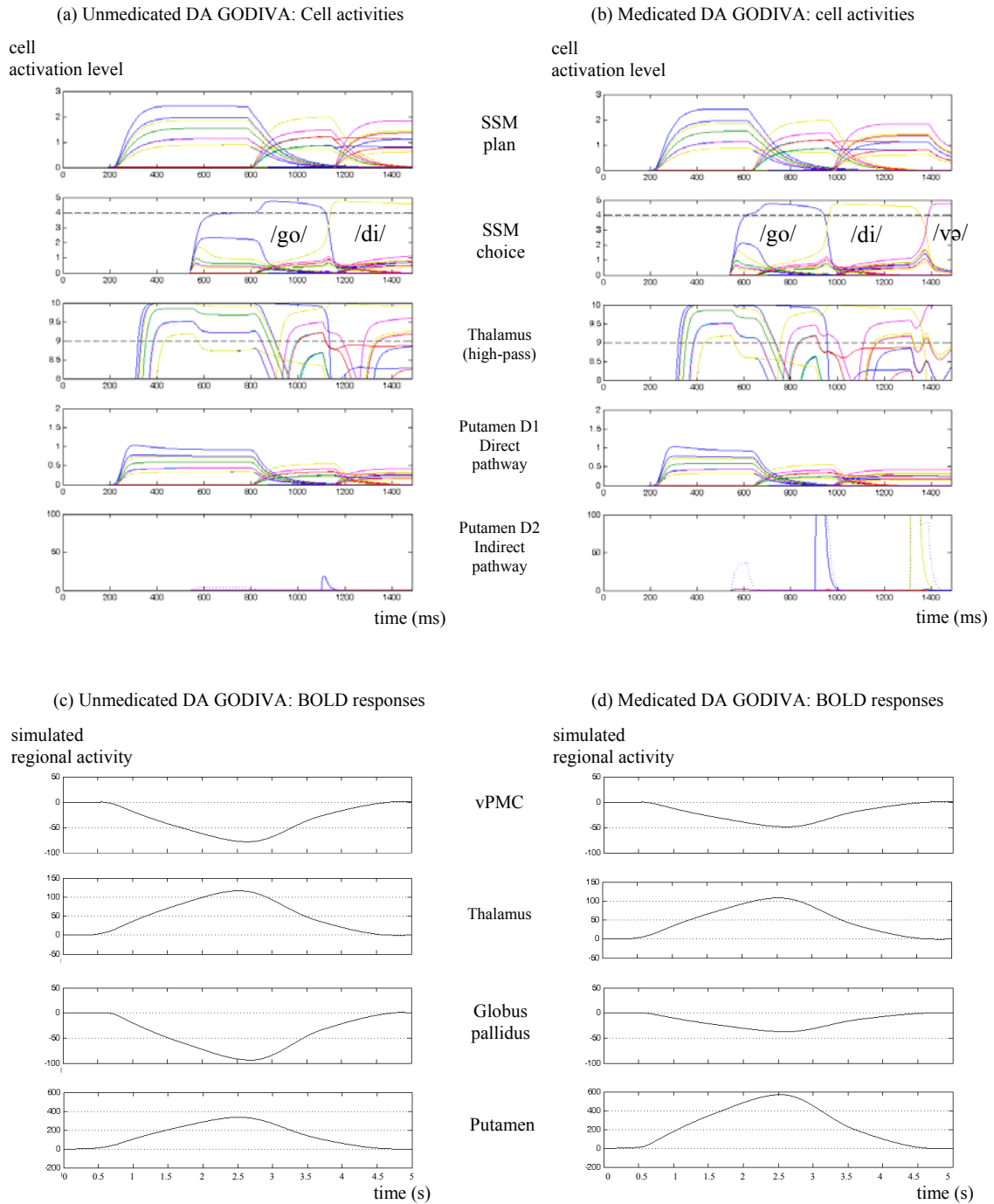
Fig. 3-4. Time course of neural activities and predicted BOLD responses of the extended GODIVA model with elevated dopamine levels. The upper panels show neural activity and follow the same conventions as in Fig. 3-3. The lower panels show the BOLD responses predicted by the simulations of the impaired model compared to the BOLD response predicted by the simulation of the intact model. (a) and (c): Unmedicated condition. (b) and (d): Under the influence of D2 antagonists.

An additional change is noted in the thalamus, where the contrast between cell activations is better than in the unmedicated condition: specifically, only the cell for "go" has maximum activation. This improvement may be attributed to the effect of blocking the D2 receptors: the putamen D2 cells increase their firing, exerting more net inhibition on the thalamus.

Fig. 3-4(c) and Fig. 3-4(d) show the BOLD responses predicted by the DA GODIVA version in unmedicated and medicated conditions, respectively, compared to the BOLD response predicted by the normal GODIVA version. For the unmedicated DA GODIVA version, the predicted BOLD responses of the putamen and thalamus are elevated, while those of the globus pallidus and vPMC are decreased. The treatment of the DA GODIVA model with D2 antagonists (medicated condition) predicts increases (relative to the unmedicated condition) in the BOLD response of all regions, except for the response of the thalamus, which decreases.

3.3.3   Dysfluent speech due to structural abnormality in white matter fibers beneath the left precentral gyrus

Fig. 3-5 shows the time course of activity in several key model components in simulations of a version of the extended GODIVA model with white-matter fibers impairment (*WMF GODIVA version*). In this version of the model, the parameter for the white matter projections' integrity, λ_{WMF}, is set to 0.025.
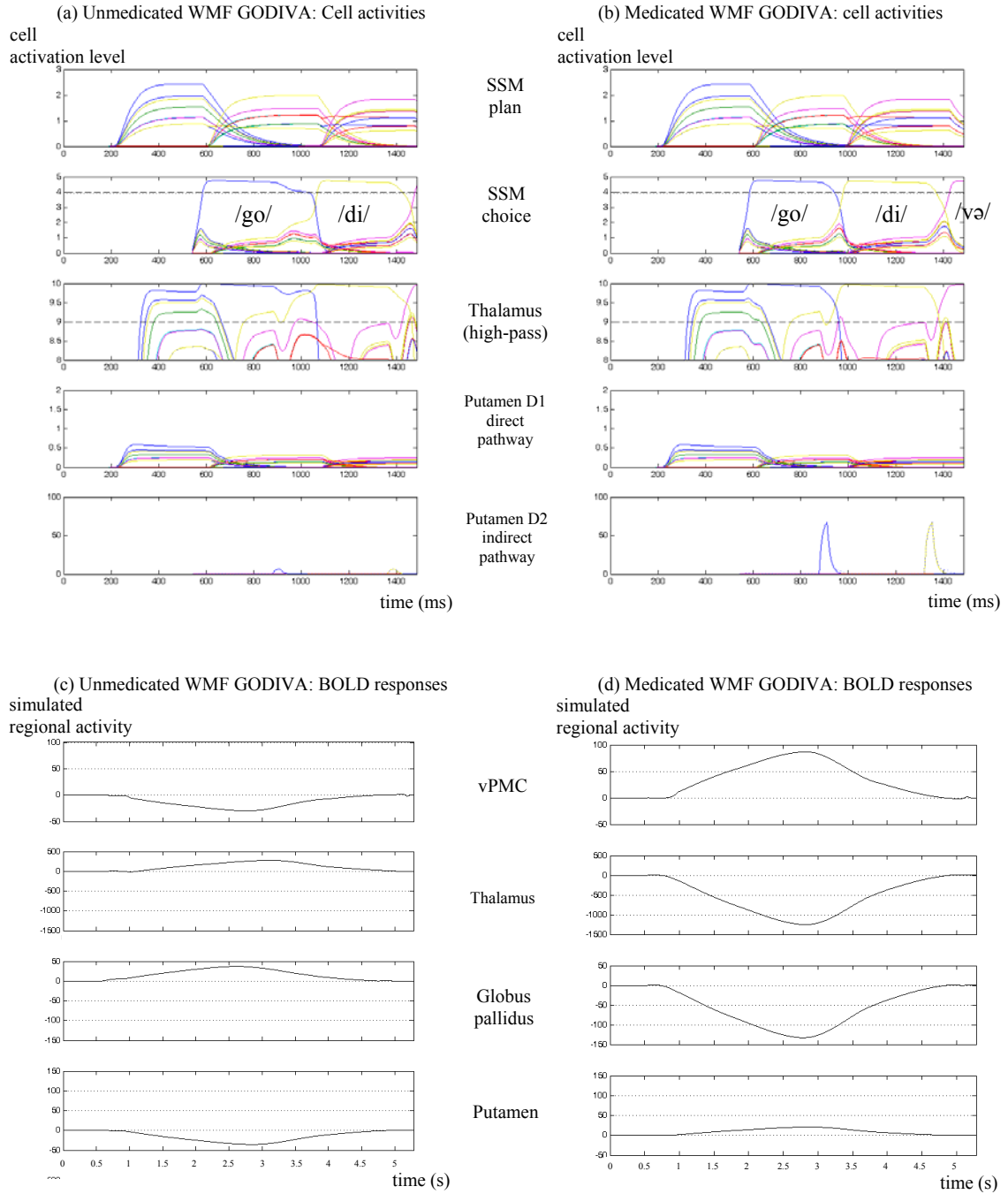
Fig. 3-5. Time course of neural activities and predicted BOLD responses of the extended GODIVA model with structural abnormality in white matter fibers. The upper panels show the neural activities and follow the same conventions as in Fig. 3-3. The lower panels show the BOLD responses predicted by the simulations of the impaired model compared to the BOLD response predicted by the simulation of the intact model. (a) and (c): Unmedicated condition. (b) and (d): Under the influence of D2 antagonists.

Fig. 3-5(a) shows a simulation of the WMF GODIVA version producing a block. A comparison to the performance of the normal GODIVA version reveals that the unmedicated WMF GODIVA version reads out the sensorimotor program of the "di" SSM too late. The winning SSM choice cell exceeds the threshold at t=1067ms, compared to t=979ms for the normal GODIVA version. The behavioral correlate is a short block (88ms long) on the second syllable of /godivə/. In this case, contrary to the DA GODIVA version, the problem is not a lack of distinction between the highest thalamic activations. Rather, it is an initial failure to generate the putamen D2 cell activity needed to quickly suppress the SSM representation of the prior syllable ("go"). Note that the same kind of block could occur on the initial syllable of a word during fluent speech, because in that case there could be failure to promptly reset the plan for the terminal syllable of the prior word.

The basis for the delayed "di" is clearly visible in Fig. 3-5(a): the transient activity in the putamen D2 cell for "di", prominent in the normal GODIVA version, is almost undetectable in the unmedicated WMF GODIVA version. The consequence of this is a weaker net indirect pathway inhibition, which can be observed as a shallower dip in the activation of the thalamus. This makes it more difficult to quench the thalamic cell supporting the SSM choice cell for the first syllable of the sequence "go.di.və". The persistence of the "go" cell as the winner in the choice layer of the vPMC prevents the selection of the "di" cell, delaying the readout of the sensorimotor program for the second syllable of the sequence.

Fig. 3-5(b) shows a simulation in which the WMF GODIVA version becomes fluent under the influence of D2 antagonists. As in the previous section, treatment with D2 antagonists was simulated as a tenfold decrease in binding of dopamine to D2 receptors ($\beta_{D1}=1$, $\beta_{D2} = 0.1$). The simulation shows that the readout of the sensorimotor program of "di" starts more quickly (t=979ms) than it does without medication, similar to the performance of the normal GODIVA version. The behavioral correlate is fluent transition between the first and second syllables of /godivə/.

In general, neural activity in the medicated WMF GODIVA version is quite similar to that in the normal GODIVA version; hence, the D2 drugs are understood to have a normalizing effect. Specifically, they elevate transient activation in the putamen D2 cells, re-enabling strong transient inhibition of the thalamus.

Fig. 3-5(c) and Fig. 3-5(d) show the BOLD responses predicted by the WMF GODIVA version in unmedicated and medicated conditions, respectively, compared to the BOLD response predicted by the normal GODIVA version. For the unmedicated WMF GODIVA version, the predicted BOLD responses of the globus pallidus and thalamus are elevated, while those of the putamen and vPMC are decreased. The simulation of the medicated WMF GODIVA version predicts a reverse pattern: the BOLD responses of the globus pallidus and thalamus are now decreased, and those of the putamen and vPMC are elevated.

3.4 Discussion

The simulations of the extended GODIVA model demonstrate two scenarios in which the SSM choice cell for the next syllable cannot be activated in a timely manner due to BG

dysfunction. In one case, the BG fail to bias cortical competition in favor of the SSM choice cell for the correct well-learned syllable, and in another, they fail to cancel the activation of the SSM choice cell for the current syllable. It is this delayed activation of SSM choice cells that leads to dysfluencies in speech, supporting Alm's (2004) claim that the BG play a major role in stuttering. The delayed activation has several possible consequences; each leads to a different type of dysfluency. As was simulated in this chapter, if no motor commands are read out at all by the SSM choice cell, the outcome is a block. If the commands to start voicing are read out, but the commands to move the articulators are not, the outcome is a prolongation[15]. If commands are improperly read out (either partial or weak readout), speech production continues, but production errors prevail. Accumulation of production errors may lead to sound/syllable repetitions, as was simulated in Chapter 2.

Furthermore, the BOLD responses predicted by the simulations of the DA GODIVA version agree with published results. It has been shown that activation levels in the putamen and thalamus are correlated with moments of stuttering (Braun et al., 1997), and that the level of activation in the globus pallidus is lower than normal in PWS (S. Brown et al., 2005). Furthermore, the simulations are consistent with repeated findings of abnormally low left vPMC activation in PWS, possibly associated with actual moments of stuttering (Ingham et al., 2004; Neumann et al., 2003; Watkins et al., 2008). The BOLD responses predicted by the WMF GODIVA version agree with some of the

---

[15] In addition to prolongations at the beginning of syllables, the model can also generate prolongations at the end of syllables. This may happen if the cortical cell that reads out the commands for a given syllable could not be quenched before the syllable is over: voicing will stay on until quenching finally occurs.

published results. The disagreement regarding the BOLD response of the putamen may suggest that the structural abnormality (in the white matter fibers) generates noise in the projections from the vPMC to the putamen. The excitations of many striatal neurons can account then for the higher activation levels of the putamen.

Dopamine has an excitatory effect on the direct pathway, and in simulations, elevated dopamine levels ultimately lead to the creation of a "ceiling effect" in the thalamus (cf. Alm, 2004). The direct pathway over-excites thalamus, pushing the activation of the thalamic cell for the correct well-learned syllable ("go") to the highest possible level; unfortunately, it does the same to cells of incorrect well-learned syllables (e.g., "gop", "god"). Although the "go" thalamic cell is still receiving stronger net-excitation than its competitors are, its activation cannot be increased above theirs because the cell has reached its ceiling. This interferes with the thalamus's role in increasing signal-to-noise ratio, or contrast, in the cortex. The signal (the excitation of the SSM choice cell for "go") is equal to the noise (the excitation of the SSM choice cells for other syllables), and does not provide the SSM choice cell for "go" with a competitive advantage. This results in slower-paced competition in the cortex, delaying the selection of the next syllable. Thus: a dysfluency.

Another conclusion drawn from the simulations is that the white matter impairment, which is assumed to affect the corticostriatal projections from the vMC to the putamen, may lead to delayed activation of the SSM choice cell for the next syllable as well. In this case, the BG's normal operation is disrupted, because the putamen D2 cells cannot reliably detect the context which indicates the imminent completion of

syllable articulation. This prevents the indirect pathway from exerting its inhibition on the thalamus, and the currently executed syllable cannot be quenched. In the current implementation of the GODIVA model, the integrity of the white matter fibers directly modulates the strength of the contextual signal generated in the putamen D2 cells, but similar results can be achieved by introducing noise into the information carried by these projections.

Finally, the simulations also account for the alleviation of stuttering with D2 antagonists (Brady, 1991). In agreement with Stager et al.'s (2005) findings, in some cases the D2 antagonists did not prevent dysfluencies, but only reduced them in duration. In both versions of the extended GODIVA model, the D2 antagonists strengthened the indirect pathway by preventing dopamine from exerting its inhibitory effect on the putamen D2 cells. In the WMF GODIVA version, the antagonists "fixed" the abnormally weak transient activation in the indirect pathway. In the DA GODIVA version, the antagonists caused the indirect pathway to inhibit the thalamus more strongly than usual, counteracting the increased excitation arriving from the direct pathway.

3.5   Conclusions

Based on GODIVA, a recently developed model of connected-speech planning and initiation, we investigate two independent hypotheses of the etiology of stuttering: the "WMF hypothesis" proposes that stuttering is caused by structural abnormalities in white matter fibers beneath the left precentral gyrus, and the "DA hypothesis" proposes that stuttering is caused by elevated dopamine levels' affecting the BG's ability to bias cortical competition in favor of an intended syllable. According to the GODIVA model

simulations, which support our predictions, either hypothesis can account for most aspects of stuttering. Here we present a summary of each hypothesis's strengths and weaknesses.

The primary weakness of the WMF hypothesis lies in its failure to account for dysfluencies on first syllables in a speech episode. Although impaired white-matter projections are hypothesized to disrupt shifts from one syllable to the next, the execution of speech-episode-initial syllables does not involve such a shift; therefore, episode-initial syllables would be spoken fluently[16]. A second flaw in the WMF hypothesis is that it does not explain the subtle non-oral motor-control deficits observed in PWS (for review, see Bloodstein & Ratner, 2008; Max, 2004; Max et al., 2004). If the white-matter impairment is indeed restricted to those corticostriatal projections that carry corollary discharge of motor commands for syllable production, the only BG loop affected should be the one that handles the selection and initiation of syllables; dysfunction in this loop would not be expected to disrupt non-oral movements. Nevertheless, the fact that the white matter abnormalities were detected at different locations in the left precentral gyrus (Chang et al., 2008; Sommer et al., 2002; Watkins et al., 2008) raises the possibility that the impairment affects several neighboring BG loops after all. Such non-specific white matter impairment could account for the non-oral deficiencies of PWS, but raises the question of why the motor behavior most affected in the stuttering disorder is speech. In

---

[16] The WMF hypothesis could account for dysfluencies on the first syllable in a sequence if we assume that motor commands for a sequence-initial syllable are preceded by pre-sequence commands, i.e., motor commands that bring the articulators from their rest position to their sequence-initial position. A failure to quench the choice cell that reads out these pre-sequence commands would then lead to dysfluency on the sequence-initial syllable.

summary, while the WMF hypothesis is supported by experimental evidence, it does not account for the full range of behaviors associated with stuttering.

The DA hypothesis, in contrast, does not lead to similar caveats. Its predicted delay in syllable selection affects all syllables, regardless of their position in the sequence. Moreover, given the non-specificity of dopamine, the implication is that all BG loops are affected by excessive levels of the neurotransmitter. This effect, however, is not universal: dopamine seems to have the greatest effect on those BG loops involved in oral motor control (Nevet, Morris, Saban, Fainstein, & Bergman, 2004). This is because the presynaptic D1 receptors in the substantia nigra pars reticulata portion of the GPi, through where the oral-motor-control BG loops traverse (Delong, Crutcher, & Georgopoulos, 1983; Gracco & Abbs, 1987), are the only receptors in the GPi that receive dopamine directly from the dendrites of the substantia nigra pars compacta (Abercrombie & Deboer, 1997; Cheramy, Nieoullon, & Glowinski, 1979). This unique property of the substantia nigra pars reticulata's presynaptic D1 receptors may be attributable to the proximity between the pars-reticulata and pars-compacta substantia nigras[17].

An additional strength of the DA hypothesis is that it implies greater normal variability across episodes of producing a given syllable (than the WMF hypothesis). Because dopamine levels fluctuate over time, the DA hypothesis can account for day-to-day variability in stuttering (Van Riper, 1982). Furthermore, because dopamine release is modulated by emotional state, the DA hypothesis can in principal account for the role of

---

[17] Another possible reason is that the SNpr has one of the densest expressions of D1 receptors in the brain (Abercrombie & Deboer, 1997).

emotions in stuttering as well (Alm, 2004**,** p. 332, 344, 359; Smith, 1999; Zimmermann, 1980b). A caveat regarding this hypothesis is the current scarcity of direct experimental evidence. We know of only one study that shows direct evidence for elevated dopamine levels in PWS (Wu et al., 1997), and while its findings were highly significant, despite its small sample size, definitive conclusions must await replications or convergent data based on further studies.

The two simulated hypotheses can be also compared in terms of their robustness, namely, whether the result of the simulations are independent of exact parameter values. According to our analysis, both hypotheses are robust because we could replicate their simulations over a wide range of parameter values (appendix C). We also evaluated both hypotheses based on the correlation between the predicted neural activities and the actual reported neural activities. The value of this approach is limited, however, by the ambiguous relationship between regional neural activities and activations of specific cell types in the GODIVA model.

Unfortunately, there is not always a one-to-one relationship between brain regions whose activities are reported in the stuttering literature and specific cell types in the GODIVA model: the putamen contains putamen D1 and putamen D2 cells (among others), the vPMC contains both SSM plan and SSM choice cells, and the globus pallidus contains both GPi and GPe cells. (Although GPi and GPe cells reside in different parts of the globus pallidus, imaging studies of stuttering have not usually differentiated between them). The thalamus does correspond to single cell type in the GODIVA model. However, because abnormal thalamic cell activity may originate from either the direct or

indirect pathways, it cannot indicate which of the pathways is dysfunctional. Given the above, we believe that reports of neural activity restricted to the GPe, the part of the globus pallidus that belongs to the indirect pathway but not the direct pathway, will prove most useful in testing the presented hypotheses.

Given its explanatory scope, its strengths in explaining some of the features of stuttering, and the close agreement between the neural activities it predicts and the imaging results, we believe the DA hypothesis warrants further investigation. We also suggest that future studies explore whether elevated dopamine levels coexist with impaired white matter projections; during brain development, the former abnormality may possibly cause the later one.

4   CONCLUSION

This dissertation presents a framework which may broaden our understanding of stuttering. Chapter 2 showed that inability to properly read out motor commands for well-learned syllables may lead to repetitions, and Chapter 3 demonstrated that neural abnormalities might be the underlying cause for the improper readout. Moreover, it showed that not only repetitions, but also pauses and prolongation, can result from the impairment.

The contributions of this work to the study of stuttering are varied. First, it presents stuttering hypotheses that both address the proximal underlying deficits (etiology), and trace the consequences of proximal deficits during actual moments of stuttering, thus meeting Bloodstein and Ratner's (2008, p. 73) criteria for "adequate" theories of stuttering. The hypotheses argue that neural abnormalities lead to impairment in the ability to read out feedforward commands, and the impairment then causes various dysfluencies, including repetitions due to overreliance on feedback control. Second, the dissertation develops neurobiologically realistic computational models of stuttering, which make it possible to test the hypotheses by comparing the model simulations to various data sets (acoustic, articulatory, and functional brain imaging data). To our knowledge, DIVA and the extended GODIVA are the first models of speech production that fit articulatory movement data collected during moments of stuttering. Finally, this dissertation is one of a handful of studies to report combined acoustic and articulatory data for moments of stuttering (for additional examples, see Harrington, 1987; van Lieshout, 2004). The combination of this multi-modal dataset and the modeling work

provides novel insights into the relation between kinematics and acoustics in stuttering; what emerges is that during repetitions, articulatory movements bring the auditory feedback closer to the desired auditory target.

4.1 Insights into the treatment of stuttering

In addition to extending our general knowledge of stuttering, this dissertation has clinical implications. The insights obtained from the simulations can help improve existing methods to treat stuttering, as well as develop new ones.

The work presented in Chapter 2 clarifies mechanisms by which some common fluency enhancers may work. This chapter utilized the DIVA model to investigate the role of auditory feedback and feedforward control in stuttering. It is hypothesized that one possible consequence of underlying neural abnormalities is a bias away from feedforward control and toward feedback control. The simulations showed that overreliance on auditory feedback may lead in some instances to sound/syllable repetitions. Furthermore, they demonstrated that this hypothesis can account for the reduction in stuttering frequency in conditions of slowed/prolonged speech or masking noise. These results may have clinical implications for popular stuttering treatment programs that use prolonged speech (Curlee, 1999) and for existing stuttering programs that use auditory feedback manipulations (Lincoln, Packman, & Onslow, 2006).

Chapter 3 may shed light on the mechanisms behind drug therapy. The chapter reports on the implementation of the GODIVA and DIVA models to investigate the role of the BG-vPMC loop in stuttering. According to the current hypothesis, both elevated dopamine levels and white matter impairment beneath the precentral gyrus can lead to

dysfunction in normal BG operation. The simulations showed that both roles of the BG; namely biasing competition in favor of the next syllable, and deleting the current syllable, can be disrupted. In both cases the result is similar: an inability to initiate the next syllable in the sequence, and hence dysfluency. By showing that the BG are the most likely brain regions to be affected by the neural abnormalities found in PWS, the GODIVA model provides an explanation for the alleviation of stuttering by D2 antagonists (Brady, 1991). In the case of elevated dopamine levels, the simulations showed that by strengthening the indirect (inhibitory) pathway of the BG-vPMC loop, D2 antagonists can counteract the hyper-activation of the direct (excitatory) pathway. This mechanistic description can accommodate the finding of increased striatal metabolism associated with stuttering (Braun et al., 1997; Giraud et al., 2007; Ingham et al., 2004; but see Wu et al., 1995) and may prove extremely useful when developing and testing dopaminergic drugs to treat the disorder (Alm, 2004; Brady, 1991; Maguire et al., 2004; Tran et al., 2008).

4.2   Future directions

To account for stuttering, a BG-vPMC loop module was added to the GODIVA model. Although the new module was extensively used throughout this dissertation, many of its functions remain unexplored. Below we discuss some future research directions prompted by this loop.

The learning that takes place in the BG is an important feature that was not addressed in the current work. The GODIVA model represents an adult speaker, and as such, no learning takes place. Instead, it is assumed that the weights between the vPMC

planning layer and the BG indirect pathway are already set such that the putamen D2 cell for the current syllable is excited when the syllable is about to complete. Modeling the process in which these weights are learned would be instructive because learning in the BG involves dopamine (see J. W. Brown et al., 2004; Frank & Claus, 2006) and is likely to be disrupted by hyper-activity in the dopaminergic system (Max, 2004; Max et al., 2004). This might provide another link between elevated dopamine levels and the BG-vPMC loop dysfunction that leads to dysfluencies.

In the current investigation the wide range of sensory inputs received by the striatum (Alexander, Crutcher, & Delong, 1990), for example, from the SMA, were deliberately overlooked. These inputs may provide the BG with more cues regarding syllable completion than those provided by the corollary discharge of executed motor commands. In normal speaking conditions sensory feedback is probably too slow to fulfill this role; however, it may prove useful in slow speech conditions (Alm & Risberg, 2009). When PWS speak slowly, the BG-vPMC loop can use the sensory information for subtle adjustments during completion of movements (cf. Max, 2004), possibly by prolonging syllables until sensory feedback indicates their completion. The modeling of the sensory inputs to the BG can potentially account for the alleviation of stuttering in slowed/prolonged speech, and unlike the explanation put forward in Chapter 2, does not presuppose sensorimotor errors.

Lastly, although the BG have complex architecture (e.g., Kopell, Rezai, Chang, & Vitek, 2006), they were here modeled as a considerably simplified circuit. One example is the absence of presynaptic dopamine receptors from the model. These receptors may

help, for example, to explain the alleviation of stuttering by low doses of dopamine agonists (Alm, 2004; Burns, Brady, & Kuruvilla, 1978). Low doses of dopamine may stimulate the striatum's presynaptic inhibitory (D2) receptors to a greater extent than the postsynaptic receptors (Brady, 1991), and the resultant weakening of the direct pathway and strengthening of the indirect pathway could normalize BG function. If we assume that stuttering is due to elevated dopamine levels, presynaptic receptors can also account for the discrepancy between recent reports of abnormal left striatal activation in PWS. The striatum of PWS may be hypo-activated (Wu et al., 1995) when presynaptic inhibitory receptors are stimulated more than postsynaptic receptors (at high dopamine levels, cf. Wolkin et al., 1987), and hyper-activated (Giraud et al., 2007) when pre- and post- synaptic receptors are stimulated to the same extent (at extremely high dopamine levels, cf. Vollenweider, Maguire, Leenders, Mathys, & Angst, 1998).

APPENDIX A. SPECIFICATION OF THE ORIGINAL GODIVA MODEL

The extended GODIVA model uses most of the equation of the original GODIVA model as-is. The only equation that was modified is the equation that describes the activities of the SSM choice cells (equation 14, given in its new form in Section 3.2.3.2). The equations that remained unchanged are listed below (adapted from Bohland et al., in press).

The activities of the IFS phonological content plan cells are given by:

$$(1) \quad \dot{p}_{ij} = -A_p p_{ij} + \left(B_p - p_{ij}\right)\left(\alpha u_{ij}^p + \left[p_{ij} - \theta_p\right]^+\right) - p_{ij}\left(\sum_{k \neq i} W_{ik} p_{kj} + 10 y\left(\left[q_{ij} - \theta_q\right]^+\right)\right) + N\left(0, \sigma_p\right)$$

The activities of the IFS phonological content choice cells are given by:

$$(2) \quad \dot{q}_{ij} = -A_q q_{ij} + \left(B_q - q_{ij}\right)\left(d_j\left[p_{ij} - \theta_p\right]^+ + y\left(q_{ij}\right)\right) - q_{ij}\left(\sum_{kj, k \neq i} W_{ik} y\left(q_{kj}\right) + \Gamma_{ij}\right)$$

The response suppression signal from vPMC choice layer to IFS choice layer is given by:

$$(3) \quad \Gamma_{ij}\left(t\right) = 10 Z_k^{ij} s_k\left(t\right)$$

The activities of the pre-SMA frame plan cells are given by:

$$(4) \quad \dot{f}_i = -A_f f_i + \left(B_f - f_i\right)\left(\alpha u_i^f + \left[f_i - \theta_f\right]^+\right) - f_i\left(\sum_{k \neq i} f_k + 10 y\left(\left[g_i - \theta_g\right]^+\right)\right)$$

The activities of the pre-SMA frame choice cells are given by:

$$(5) \quad \dot{g}_i = -A_g g_i + \left(B_g - g_i\right)\left(\Omega\left[f_i - \theta_f\right]^+ + y\left(g_i\right)\right) - g_i\left(\sum_{k \neq i} y\left(g_k\right)\right)$$

The activities of the pre-SMA positional chain cells are given by:

$$(6) \quad h_j^k\left(t\right) = \begin{cases} 1 & \text{if } \left(t_0 + (j-1)\tau\right) \leq t \leq \left(t_0 + j\tau\right) \\ 0 & \text{otherwise} \end{cases}$$

The activities of the planning loop striatal projection neurons (direct pathway) are given by:

$$(7) \quad \dot{b}_j = -A_b b_j + \left(B_b - b_j\right)\left(h_j \wedge \left[\sum_k p_{kj} - \delta\right]^+\right) - b_j\left(\sum_{k \neq j} y(\underline{b}_k)\right)$$

The activities of the planning loop striatal interneurons are given by:

$$(8) \quad \dot{\underline{b}}_j = -A_{\underline{b}} \underline{b}_j + \left(B_{\underline{b}} - \underline{b}_j\right)\left(h_j \wedge \left[\sum_k p_{kj} - \delta\right]^+\right) - \underline{b}\left(\sum_{k \neq j} y(\underline{b}_k)\right)$$

The activities of the planning loop GPi cells are given by:

$$(9) \quad \dot{c}_j = -A_c c_j + \beta_c\left(B_c - c_j\right) - c_j\left(b_j\right)$$

The activities of the planning loop anterior thalamic cells are given by:

$$(10) \quad \dot{d}_j = -A_d d_j + \beta_d\left(B_d - d_j\right) - d_j\left(c_j\right)$$

The synapse from IFS choice cells to SSM plan cells that code multi-phoneme targets are given by:

$$(11) \quad Z_k^{ij} = \begin{cases} \dfrac{1}{N_k} & \text{if } r_k \text{ includes phoneme } i \text{ at position } j \\ 0 & \text{otherwise} \end{cases}$$

The synapses from IFS choice cells to SSM plan cells that code single phoneme targets are given by:

$$(12) \quad Z_k^{ij} = \begin{cases} 0.85 - 0.05j & \text{if } r_k \text{ codes phoneme } i \\ 0 & \text{otherwise} \end{cases}$$

The activities of the SSM plan cells are given by:

$$(13) \quad \dot{r}_k = -A_r r_k + \left(B_r - r_k\right)\left(\sum_i \sum_j Z_k^{ij} y\left(\left[q_{ij} - \theta_q\right]^+\right) + \left[r_k - \theta_r\right]^+\right) - r_k\left(\sum_{n \neq k} r_n\right)$$

APPENDIX B. VARIABLES OF THE ORIGINAL GODIVA MODEL

Table B-1. Legend of main symbols used to refer to cell populations in the specification of the original GODIVA model (reproduced from Bohland et al., in press).

Cell Group	Symbol
External Input to IFS	u^p
External Input to preSMA	u^f
IFS Phonological Content Plan Cells	p
IFS Phonological Content Choice Cells	q
Pre-SMA Frame Plan Cells	f
Pre-SMA Frame Choice Cells	g
Pre-SMA Positional Chain Cells	h
Planning Loop Striatal Projection Cells	b
Planning Loop Striatal Interneurons	$\underline{b}$
Planning Loop GPi Cells	c
Planning Loop Anterior Thalamic Cells	d
SSM Plan Cells (vPMC planning layer)	r

APPENDIX C. CRITICAL PARAMETER VALUE RANGES

Table C-1. Critical parameter value ranges for normal and abnormal versions of the extended GODIVA model.

Description	Notation	Normal values (persons who not stutter)		Abnormal values (persons who stutter)	
		Simulated	Range	Simulated	Range
Bias toward feedforward control	α_{ff}	0.85	0.46-1.0	0.25	0-0.46
Bias toward feedback control	α_{fb}	0.15	0-0.54	0.75	0.54-1.0
Binding of dopamine to D1 receptors	β_{D1}	1.0	0.5-1.5	1.75	1.5-3.0
Binding of dopamine to D2 receptors	β_{D2}	1.0	0.5-1.5	1.75	1.5-3.0
Integrity of white matter fibers	λ_{WMF}	1.0	0.04-3	0.025	0-0.04

REFERENCES

Abbs, J. H. (1973). The influence of the gamma motor system on jaw movements during speech: A theoretical framework and some preliminary observations. *Journal of Speech and Hearing Research, 16*(2), 175-200.

Abercrombie, E. D., & Deboer, P. (1997). Substantia nigra D1 receptors and stimulation of striatal cholinergic interneurons by dopamine: A proposed circuit mechanism. *Journal of Neuroscience, 17*(21), 8498-8505.

Adams, M. R. (1978). Further analysis of stuttering as a phonetic transition defect. *Journal of Fluency Disorders, 3*, 265-271.

Adams, M. R., & Hutchinson, J. (1974). The effects of three levels of auditory masking on selected vocal characteristics and the frequency of disfluency of adult stutterers. *Journal of Speech and Hearing Research, 17*(4), 682-688.

Adams, M. R., & Moore, W. H., Jr. (1972). The effects of auditory masking on the anxiety level, frequency of dysfluency, and selected vocal characteristics of stutterers. *Journal of Speech and Hearing Research, 15*(3), 572-578.

Agnello, J. G. (1975). Voice onset and voice termination features of stutterers. In L. M. Webster & L. C. Furst (Eds.), *The first annual Hayes Martin Conference on Vocal Tract Dynamics. Vocal tract dynamics and dysfluency* (pp. 40-70). New York: Speech and Hearing Institute.

Albin, R. L., Young, A. B., & Penney, J. B. (1989). The functional anatomy of basal ganglia disorders. *Trends in Neurosciences, 12*, 366-374.

Alexander, G. E., Crutcher, M. D., & Delong, M. R. (1990). Basal ganglia-thalamocortical circuits: Parallel substrates for motor, oculomotor, 'prefrontal' and 'limbic' functions. *Progress in Brain Research, 85*, 119-146.

Alfonso, P. J. (1991). Implications of the concepts underlying task-dynamic modeling on kinematic studies of stuttering. In H. F. M. Peters, W. Hulstijn & C. W. Starkweather (Eds.), *International Conference on Speech Motor Control and Stuttering (2nd : 1990 : Nijmegen, Netherlands). Speech motor control and stuttering* (pp. 79-100). Amsterdam: Excerpta Medica.

Alm, P. A. (2004). Stuttering and the basal ganglia circuits: A critical review of possible relations. *Journal of Communication Disorders, 37*(4), 325-369.

Alm, P. A., & Risberg, J. (2009). *The role of auditory feedback in persons who stutter: A study of vowel duration control*. Manuscript submitted for publication.

Altrows, I. F., & Bryden, M. P. (1977). Temporal factors in the effects of masking noise on fluency of stutterers. *Journal of Communication Disorders, 10*(4), 315-329.

Andrews, G., Craig, A., Feyer, A. M., Hoddinott, S., Howie, P., & Neilson, M. (1983). Stuttering: A review of research findings and theories circa 1982. *Journal of Speech and Hearing Disorders, 48*(3), 226-246.

Andrews, G., Howie, P. M., Dozsa, M., & Guitar, B. E. (1982). Stuttering: Speech pattern characteristics under fluency-inducing conditions. *Journal of Speech and Hearing Research, 25*(2), 208-216.

Armson, J., & Kalinowski, J. (1994). Interpreting results of the fluent speech paradigm in stuttering research: Difficulties in separating cause from effect. *Journal of Speech and Hearing Research, 37*(1), 69-82.

Barbas, H., & Pandya, D. N. (1987). Architecture and frontal cortical connections of the premotor cortex (area 6) in the rhesus monkey. *Journal of Comparative Neurology, 256*(2), 211-228.

Barr, D. F., & Carmel, N. R. (1969). Stuttering inhibition with high frequency narrow-band masking noise. *The journal of Auditory Research, 9*, 40-44.

Baum, S. R., McFarland, D. H., & Diab, M. (1996). Compensation to articulatory perturbation: Perceptual data. *Journal of the Acoustical Society of America, 99*(6), 3791-3794.

Blomgren, M., Robb, M., & Chen, Y. (1998). A note on vowel centralization in stuttering and nonstuttering individuals. *Journal of Speech, Language, and Hearing Research, 41*, 1042-1051.

Bloodstein, O. (1995). *A handbook on stuttering* (5th ed.). San Diego, CA: Singular Publishing Group.

Bloodstein, O., & Ratner, N. B. (2008). *A handbook on stuttering* (6th ed.). Clifton Park, NY: Thomson/Delmar Learning.

Bohland, J. W. (2007). Neuroimaging and computational modeling of syllable sequence production (Doctoral dissertation, Boston University, 2007). *Dissertation Abstracts International*, *68*, 03.

Bohland, J. W., Bullock, D. H., & Guenther, F. H. (in press). Neural representations and mechanisms for the performance of simple speech sequences. *Journal of Cognitive Neuroscience*.

Bohland, J. W., & Guenther, F. H. (2006). An fMRI investigation of syllable sequence production. *Neuroimage, 32*(2), 821-841.

Borden, G. J. (1979). An interpretation of research on feedback interruption in speech. *Brain and Language, 7*, 307-319.

Brady, J. P. (1991). The pharmacology of stuttering: A critical review. *American Journal of Psychiatry, 148*, 1309-1316.

Braun, A. R., Varga, M., Stager, S., Schulz, G., Selbie, S., Maisog, J. M., et al. (1997). Altered patterns of cerebral activity during speech and language production in developmental stuttering. An H2(15)O positron emission tomography study. *Brain, 120*(Pt 5), 761-784.

Brotchie, P., Iansek, R., & Horne, M. K. (1991). Motor function of the monkey globus pallidus. 2. Cognitive aspects of movement and phasic neuronal activity. *Brain, 114*(Pt 4), 1685-1702.

Brown, J. W., Bullock, D., & Grossberg, S. (2004). How laminar frontal cortex and basal ganglia circuits interact to control planned and reactive saccades. *Neural Networks, 17*(4), 471-510.

Brown, S., Ingham, R. J., Ingham, J. C., Laird, A. R., & Fox, P. T. (2005). Stuttered and fluent speech production: An ALE meta-analysis of functional neuroimaging studies. *Human Brain Mapping, 25*(1), 105-117.

Brown, S. F. (1938). A further study of stuttering in relation to various speech sounds. *Quarterly Journal of Speech, 24*, 390-397.

Brown, T., Sambrooks, J. E., & MacCulloch, M. J. (1975). Auditory thresholds and the effect of reduced auditory feedback on stuttering. *Acta Psychiatrica Scandinavica, 51*, 297-311.

Burger, R., & Wijnen, F. (1999). Phonological encoding and word stress in stuttering and nonstuttering subjects. *Journal of Fluency Disorders, 24*(2), 91-106.

Burns, D., Brady, J. P., & Kuruvilla, K. (1978). The acute effect of haloperidol and apomorphine on the severity of stuttering. *Biological Psychiatry, 13*(2), 255-264.

Callan, D. E., Kent, R. D., Guenther, F. H., & Vorperian, H. K. (2000). An auditory-feedback-based neural network model of speech production that is robust to developmental changes in the size and shape of the articulatory system. *Journal of Speech, Language, and Hearing Research, 43*(3), 721-736.

Caruso, A. J., Abbs, J. H., & Gracco, V. L. (1988). Kinematic analysis of multiple movement coordination during speech in stutterers. *Brain, 111*(Pt 2), 439-456.

Caruso, A. J., Chodzko-Zajko, W. J., Bidinger, D. A., & Sommers, R. K. (1994). Adults who stutter: Responses to cognitive stress. *Journal of Speech and Hearing Research, 37*(4), 746-754.

Chang, S. E., Erickson, K. I., Ambrose, N. G., Hasegawa-Johnson, M. A., & Ludlow, C. L. (2008). Brain anatomy differences in childhood stuttering. *Neuroimage, 39*(3), 1333-1344.

Chang, S. E., Ohde, R. N., & Conture, E. G. (2002). Coarticulation and formant transition rate in young children who stutter. *Journal of Speech, Language, and Hearing Research, 45*(4), 676-688.

Chase, R. A., & Sutton, S. (1968). Reply to: "masking of auditory feedback in stutterers' speech". *Journal of Speech and Hearing Research, 11*, 222-223.

Cheramy, A., Nieoullon, A., & Glowinski, J. (1979). In vivo evidence for a dendritic release of dopamine in the cat substantia nigra. *Applied Neurophysiology, 42*(1-2), 57-59.

Cherry, C., & Sayers, B. M. (1956). Experiments upon the total inhibition of stammering by external control, and some clinical results. *Journal of Psychosomatic Research, 1*(4), 233-246.

Civier, O., & Guenther, F. H. (2005, June-July). *Simulations of feedback and feedforward control in stuttering.* Paper presented at the 7th Oxford Dysfluency Conference, St. Catherine's College, Oxford University.

Conture, E. G. (1991). Young stutterers' speech production: A critical review. In H. F. M. Peters, W. Hulstijn & C. W. Starkweather (Eds.), *International Conference on Speech Motor Control and Stuttering (2nd : 1990 : Nijmegen, Netherlands). Speech motor control and stuttering* (pp. 365-384). Amsterdam: Excerpta Medica.

Conture, E. G. (2001). *Stuttering: Its nature, diagnosis, and treatment*. Boston: Allyn and Bacon.

Conture, E. G., & Brayton, E. R. (1975). The influence of noise on stutterers' different disfluency types. *Journal of Speech and Hearing Research, 18*, 381-384.

Conture, E. G., Zackheim, C. T., Anderson, J. D., & Pellowski, M. W. (2004). Linguistic processes and childhood stuttering: Many's a slip between intention and lip. In B. Maassen, R. Kent, H. F. M. Peters, P. H. H. M. van Lieshout & W. Hulstijn (Eds.). *Speech motor control in normal and disordered speech* (pp. 253-281). Oxford, UK: Oxford University Press.

Cowan, R. L., & Wilson, C. J. (1994). Spontaneous firing patterns and axonal projections of single corticostriatal neurons in the rat medial agranular cortex. *Journal of Neurophysiology, 71*(1), 17-32.

Cowie, R. I., & Douglas-Cowie, E. (1983). Speech production in profound post-lingual deafness. In M. E. Lutman & M. P. Haggard (Eds.). *Hearing science and hearing disorders* (pp. 183-231). New York: Academic Press.

Crosson, B. (1992). *Subcortical functions in language and memory*. New York: Guilford.

Curlee, R. F. (1999). *Stuttering and related disorders of fluency* (2. ed.). New York: Thieme Medical Publishers.

Davidow, J. H., Bothe, A. K., Andreatta, R. D., & Ye, J. (2008). Measurement of phonated intervals during four fluency-inducing conditions. *Journal of Speech, Language, and Hearing Research*.

De Nil, L. F. (2004). Recent developments in brain imaging research in stuttering. In B. Maassen, R. Kent, H. F. M. Peters, P. van Lieshout & W. Hulstijn (Eds.). *Speech motor control in normal and disordered speech* (pp. 113-137). Oxford, England: Oxford University Press.

De Nil, L. F., Kroll, R. M., & Houle, S. (2001). Functional neuroimaging of cerebellar activation during single word reading and verb generation in stuttering and nonstuttering adults. *Neuroscience Letters, 302*(2-3), 77-80.

Delong, M. R. (1971). Activity of pallidal neurons during movement. *Journal of Neurophysiology, 34*(3), 414-427.

Delong, M. R., Crutcher, M. D., & Georgopoulos, A. P. (1983). Relations between movement and single cell discharge in the substantia nigra of the behaving monkey. *Journal of Neuroscience, 3*(8), 1599-1606.

Deniau, J. M., Menetrey, A., & Charpier, S. (1996). The lamellar organization of the rat substantia nigra pars reticulata: Segregated patterns of striatal afferents and

relationship to the topography of corticostriatal projections. *Neuroscience, 73*(3), 761-781.

Fairbanks, G. (1954). Systematic research in experimental phonetics: 1. A theory of the speech mechanism as a servosystem. *Journal of Speech and Hearing Disorders, 19*(2), 133-139.

Filatova, E. V., Orlov, A. A., Tolkunov, B. F., & Afanas'ev, S. V. (2005). Neuron activity in the monkey striatum of identifies integration sequential actions into functional blocks. *Neuroscience and Behavioral Physiology, 35*(9), 943-949.

Flaherty, A. W., & Graybiel, A. M. (1994). Input-output organization of the sensorimotor striatum in the squirrel monkey. *Journal of Neuroscience, 14*(2), 599-610.

Frank, M. J. (2005). Dynamic dopamine modulation in the basal ganglia: A neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. *Journal of Cognitive Neuroscience, 17*, 51-72.

Frank, M. J., & Claus, E. D. (2006). Anatomy of a decision: Striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychological Review, 113*(2), 300-326.

Gammon, S. A., Smith, P. J., Daniloff, R. G., & Kim, C. W. (1971). Articulation and stress-juncture production under oral anesthetization and masking. *Journal of Speech and Hearing Research, 14*(2), 271-282.

Garber, S. F., & Martin, R. R. (1977). Effects of noise and increased vocal intensity on stuttering. *Journal of Speech and Hearing Research, 20*(2), 233-240.

Gerfen, C. R. (1992). The neostriatal mosaic: Multiple levels of compartmental organization in the basal ganglia. *Annual Review of Neuroscience, 15*, 285-320.

Ghosh, S. S. (2005). Understanding cortical and cerebellar contributions to speech production through modeling and functional imaging (Doctoral dissertation, Boston University, 2005). *Dissertation Abstracts International, 66*, 11.

Ghosh, S. S., Tourville, J. A., & Guenther, F. H. (2008). A neuroimaging study of premotor lateralization and cerebellar involvement in the production of phonemes and syllables. *Journal of Speech, Language, and Hearing Research, 51*(5), 1183-1202.

Giraud, A. L., Neumann, K., Bachoud-Levi, A. C., Von Gudenberg, A. W., Euler, H. A., Lanfermann, H., et al. (2007). Severity of dysfluency correlates with basal ganglia activity in persistent developmental stuttering. *Brain and Language*.

Goehl, H., & Kaufman, D. K. (1984). Do the effects of adventitious deafness include disordered speech? *Journal of Speech and Hearing Disorders, 49*(1), 58-64.

Gracco, V. L., & Abbs, J. H. (1987). Programming and execution processes of speech movement control: Potential neural correlates. In E. Keller & M. Gopnik (Eds.). *Motor and sensory processes of language* (pp. xix, 306 p.). Hillsdale, N.J.: L. Erlbaum Associates.

Grossberg, S. (1973). Contour enhancement, short term memory, and constancies in reverberating neural networks. *Studies in Applied Mathematics, 52*(3), 213–257.

Grossberg, S. (1978). A theory of human memory: Self-organization and performance of sensory-motor codes, maps, and plans. In R. Rosen & F. Snell (Eds.). *Progress in theoretical biology* (Vol. 5, pp. 233-374). New York: Academic Press.

Guenther, F. H. (1994). A neural network model of speech acquisition and motor equivalent speech production. *Biological Cybernetics, 72*(1), 43-53.

Guenther, F. H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review, 102*(3), 594-621.

Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *Journal of Communication Disorders, 39*(5), 350-365.

Guenther, F. H. (2008). Neuroimaging of normal speech production. In R. J. Ingham (Ed.). *Neuroimaging in communication sciences and disorders*. San Diego, CA: Plural Pub.

Guenther, F. H., Espy-Wilson, C. Y., Boyce, S. E., Matthies, M. L., Zandipour, M., & Perkell, J. S. (1999). Articulatory tradeoffs reduce acoustic variability during American English /r/ production. *Journal of the Acoustical Society of America, 105*(5), 2854-2865.

Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language, 96*(3), 280-301.

Guenther, F. H., Hampson, M., & Johnson, D. (1998). A theoretical investigation of reference frames for the planning of speech movements. *Psychological Review, 105*(4), 611-633.

Gurney, K., Prescott, T. J., & Redgrave, P. (2001a). A computational model of action selection in the basal ganglia. I. A new functional anatomy. *Biological Cybernetics, 84*(6), 401-410.

Gurney, K., Prescott, T. J., & Redgrave, P. (2001b). A computational model of action selection in the basal ganglia. II. Analysis and simulation of behaviour. *Biological Cybernetics, 84*(6), 411-423.

Harrington, J. (1987). Coarticulation and stuttering: An acoustic and electropalatographic study. In H. F. M. Peters & W. Hulstijn (Eds.), *Speech Motor Dynamics in Stuttering (1985 : University of Nijmegen). Speech motor dynamics in stuttering* (pp. 381-392). New York: Springer-Verlag.

Harrington, J. (1988). Stuttering, delayed auditory feedback, and linguistic rhythm. *Journal of Speech and Hearing Research, 31*(1), 36-47.

Healey, E. C., & Adams, M. R. (1981). Rate reduction strategies used by normally fluent and stuttering children and adults. *Journal of Fluency Disorders, 6*, 1-14.

Healey, E. C., & Gutkin, B. (1984). Analysis of stutterers' voice onset times and fundamental frequency contours during fluency. *Journal of Speech and Hearing Research, 27*(2), 219-225.

Heuer, R. J., Sataloff, R. T., Mandel, S., & Travers, N. (1996). Neurogenic stuttering: Further corroboration of site of lesion. *Ear, Nose, and Throat Journal, 75*(3), 161-168.

Hikosaka, O., Takikawa, Y., & Kawagoe, R. (2000). Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiological Reviews, 80*(3), 953-978.

Hirsch, F., Fauvet, F., Ferbach-Hecker, V., Béchet, M., Bouarourou, F., & Sturm, J. (2007). Formant structures of vowels produced by stutterers at normal and fast speech rates. In J. Trouvain (Ed.), *16th International Congress of Phonetic Sciences, Saarbrücken, Germany* (pp. 1345-1348). Saarbrücken: Universität des Saarlandes.

Hoole, P. (1987). Bite-block speech in the absence of oral sensibility. In K. Bolla (Ed.), *11th International Congress on Phonetic Sciences Tallin* (Vol. 4, pp. 16-19). Budapest, Hungary: Academy of Sciences of the Estonian SSR.

Hoover, J. E., & Strick, P. L. (1993). Multiple output channels in the basal ganglia. *Science, 259*(5096), 819-821.

Howell, P., & Vause, L. (1986). Acoustic analysis and perception of vowels in stuttered speech. *Journal of the Acoustical Society of America, 79*(5), 1571-1579.

Howell, P., Williams, M., & Vause, L. (1987). Acoustic analysis of repetitions in stutterers' speech. In H. F. M. Peters & W. Hulstijn (Eds.), *Speech Motor Dynamics in Stuttering (1985 : University of Nijmegen). Speech motor dynamics in stuttering* (pp. 371-380). New York: Springer-Verlag.

Hutchinson, J. M., & Norris, G. M. (1977). The differential effect of three auditory stimuli on the frequency of stuttering behaviors. *Journal of Fluency Disorders, 2*, 283-293.

Hutchinson, J. M., & Ringel, R. L. (1975). The effect of oral sensory deprivation on stuttering behavior. *Journal of Communication Disorders, 8*(3), 249-258.

Ingham, R. J. (2001). Brain imaging studies of developmental stuttering. *Journal of Communication Disorders, 34*(6), 493-516.

Ingham, R. J. (2008). Neuroimaging contributions to developmental stuttering theory and treatment. In R. J. Ingham (Ed.). *Neuroimaging in communication sciences and disorders*. San Diego, CA: Plural Pub.

Ingham, R. J., Fox, P. T., Ingham, J. C., Xiong, J., Zamarripa, F., Hardies, L. J., et al. (2004). Brain correlates of stuttering and syllable production: Gender comparison and replication. *Journal of Speech, Language, and Hearing Research, 47*(2), 321-341.

Ingham, R. J., Ingham, J. C., Finn, P., & Fox, P. T. (2003). Towards a functional neural systems model of developmental stuttering. *Journal of Fluency Disorders, 28*(4), 297-317.

Ingham, R. J., Kilgo, M., Ingham, J. C., Moglia, R., Belknap, H., & Sanchez, T. (2001). Evaluation of a stuttering treatment based on reduction of short phonation intervals. *Journal of Speech, Language, and Hearing Research, 44*(6), 1229-1244.

Ingham, R. J., Martin, R. R., & Kuhl, P. (1974). Modification and control of rate of speaking by stutterers. *Journal of Speech and Hearing Research, 17*(3), 489-496.

Jäncke, L. (1991). The 'audio-phonatoric' coupling in stuttering and nonstuttering adults: Experimental contributions. In H. F. M. Peters, W. Hulstijn & C. W. Starkweather (Eds.), *International Conference on Speech Motor Control and Stuttering (2nd : 1990 : Nijmegen, Netherlands). Speech motor control and stuttering* (pp. 171-180). Amsterdam: Excerpta Medica.

Kalveram, K. T. (1991). How pathological audio-phonatoric coupling induces stuttering: A model of speech flow control. In H. F. M. Peters, W. Hulstijn & C. W. Starkweather (Eds.), *International Conference on Speech Motor Control and Stuttering (2nd : 1990 : Nijmegen, Netherlands). Speech motor control and stuttering* (pp. 163-169). Amsterdam: Excerpta Medica.

Kalveram, K. T. (1993). A neural-network model enabling sensorimotor learning: Application to the control of arm movements and some implications for speech-motor control and stuttering. *Psychological Research. Psychologiche Forschung, 55*, 299-314.

Kalveram, K. T., & Jäncke, L. (1989). Vowel duration and voice onset time for stressed and nonstressed syllables in stutterers under delayed auditory feedback condition. *Folia Phoniatrica, 41*(1), 30-42.

Kent, R. D. (1976). Models of speech production. In N. J. Lass (Ed.). *Contemporary issues in experimental phonetics* (pp. 79-104). New York: Academic Press.

Kent, R. D., & Moll, K. L. (1975). Articulatory timing in selected consonant sequences. *Brain and Language, 2*, 304-323.

Kitai, S. T., Sugimori, M., & Kocsis, J. D. (1976). Excitatory nature of dopamine in the nigro-caudate pathway. *Experimental Brain Research, 24*(4), 351-363.

Klich, R. J., & May, G. M. (1982). Spectrographic study of vowels in stutterers' fluent speech. *Journal of Speech and Hearing Research, 25*(3), 364-370.

Kolk, H., & Postma, A. (1997). Stuttering as a covert repair phenomenon. In R. F. Curlee & G. M. Siegel (Eds.). *Nature and treatment of stuttering: New directions* (2nd ed., pp. 182-203). Boston: Allyn and Bacon.

Koller, W. C. (1983). Dysfluency (stuttering) in extrapyramidal disease. *Archives of Neurology, 40*(3), 175-177.

Koos, T., & Tepper, J. M. (1999). Inhibitory control of neostriatal projection neurons by gabaergic interneurons. *Nature Neuroscience, 2*(5), 467-472.

Kopell, B. H., Rezai, A. R., Chang, J. W., & Vitek, J. L. (2006). Anatomy and physiology of the basal ganglia: Implications for deep brain stimulation for Parkinson's disease. *Movement Disorders, 21 Suppl 14*, S238-246.

Kunzle, H. (1978). An autoradiographic analysis of the efferent connections from premotor and adjacent prefrontal regions (areas 6 and 9) in macaca fascicularis. *Brain, Behavior and Evolution, 15*(3), 185-234.

Lane, H., & Tranel, B. (1971). The Lombard sign and the role of hearing in speech. *Journal of Speech, Language, and Hearing Research, 14*, 677-709.

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.). *Cerebral mechanisms in behavior* (pp. 507-528). New York: Wiley.

Lei, W., Jiao, Y., Del Mar, N., & Reiner, A. (2004). Evidence for differential cortical input to direct pathway versus indirect pathway striatal projection neurons in rats. *Journal of Neuroscience, 24*(38), 8289-8299.

Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition, 14*(1), 41-104.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Lieberman, P., & Blumstein, S. (1988). *Speech physiology, speech perception, and acoustic phonetics*. Cambridge, UK: Cambridge University Press.

Lincoln, M., Packman, A., & Onslow, M. (2006). Altered auditory feedback and the treatment of stuttering: A review. *Journal of Fluency Disorders, 31*(2), 71-89.

Liu, C., & Kewley-Port, D. (2004). Formant discrimination in noise for isolated vowels. *Journal of the Acoustical Society of America, 116*(5), 3119-3129.

Loucks, T. M., & De Nil, L. F. (2006a). Anomalous sensorimotor integration in adults who stutter: A tendon vibration study. *Neuroscience Letters, 402*(1-2), 195-200.

Loucks, T. M., & De Nil, L. F. (2006b). Oral kinesthetic deficit in adults who stutter: A target-accuracy study. *Journal of motor behavior, 38*(3), 238-246.

Ludlow, C. L., & Loucks, T. (2003). Stuttering: A dynamic motor control disorder. *Journal of Fluency Disorders, 28*(4), 273-295.

Ludlow, C. L., Rosenberg, J., Salazar, A., Grafman, J., & Smutok, M. (1987). Site of penetrating brain lesions causing chronic acquired stuttering. *Annals of Neurology, 22*(1), 60-66.

MacCulloch, M. J., Eaton, R., & Long, E. (1970). The long term effect of auditory masking on young stutterers. *The British Journal of Disorders of Communication, 5*(2), 165-173.

Macleod, J., Kalinowski, J., Stuart, A., & Armson, J. (1995). Effect of single and combined altered auditory feedback on stuttering frequency at two speech rates. *Journal of Communication Disorders, 28*(3), 217-228.

Maeda, S. (1990). Compensatory articulation during speech: Evidence from the analysis and synthesis of vocal tract shapes using an articulatory model. In W. J. Hardcastle & A. Marchal (Eds.). *Speech production and speech modelling* (pp. 131-149). Boston: Kluwer Academic Publishers.

Maguire, G. A., & Snyder, G. (2008, August 18). Dr. Gerald Maguire: Pharmaceutical treatment and research (episode 51). Interview posted to http://stuttertalk.com/2008/08/18/gerald-maguire-episode-51.Aspx

Maguire, G. A., Yu, B. P., Franklin, D. L., & Riley, G. D. (2004). Alleviating stuttering with pharmacological interventions. *Expert Opinion on Pharmacotherapy, 5*(7), 1565-1571.

Maraist, J. A., & Hutton, C. (1957). Effects of auditory masking upon the speech of stutterers. *Journal of Speech and Hearing Disorders, 22*(3), 385-389.

Martin, J. E. (1970). The signal detection hypothesis and perceptual defect theory of stuttering. *Journal of Speech and Hearing Research, 35*, 352-355.

Max, L. (2004). Stuttering and internal models for sensorimotor control: A theoretical perspective to generate testable hypotheses. In B. Maassen, R. Kent, H. F. M. Peters, P. H. H. M. van Lieshout & W. Hulstijn (Eds.). *Speech motor control in normal and disordered speech* (pp. 357–387). Oxford, UK: Oxford University Press.

Max, L., Guenther, F. H., Gracco, V. L., Ghosh, S. S., & Wallace, M. E. (2004). Unstable or insufficiently activated internal models and feedback-biased motor control as sources of dysfluency: A theoretical model of stuttering. *Contemporary Issues in Communication Science and Disorders, 31*, 105-122.

McClay, H., & Osgood, C. E. (1959). Hesitation phenomena in spontaneous English speech. *Word, 15*, 19-44.

McClean, M. D., & Runyan, C. M. (2000). Variations in the relative speeds of orofacial structures with stuttering severity. *Journal of Speech, Language, and Hearing Research, 43*(6), 1524-1531.

McClean, M. D., Tasko, S. M., & Runyan, C. M. (2004). Orofacial movements associated with fluent speech in persons who stutter. *Journal of Speech, Language, and Hearing Research, 47*(2), 294-303.

Melnick, K. S., & Conture, E. G. (2000). Relationship of length and grammatical complexity to the systematic and nonsystematic speech errors and stuttering of children who stutter. *Journal of Fluency Disorders, 25*(1), 21-45.

Melnick, K. S., Conture, E. G., & Ohde, R. N. (2003). Phonological priming in picture naming of young children who stutter. *Journal of Speech, Language, and Hearing Research, 46*(6), 1428-1443.

Miller, J. L., & Baer, T. (1983). Some effects of speaking rate on the production of /b/ and /w/. *Journal of the Acoustical Society of America, 73*(5), 1751-1755.

Mink, J. W. (1996). The basal ganglia: Focused selection and inhibition of competing motor programs. *Progress in Neurobiology, 50*(4), 381-425.

Mink, J. W., & Thach, W. T. (1993). Basal ganglia intrinsic circuits and their role in behavior. *Current Opinion in Neurobiology, 3*(6), 950-957.

Montgomery, B. M., & Fitch, J. L. (1988). The prevalence of stuttering in the hearing-impaired school age population. *Journal of Speech and Hearing Disorders, 53*(2), 131-135.

Moore, B. C. J. (1995). *An introduction to the psychology of hearing* (3rd ed.). London: Academic Press.

Muakkassa, K. F., & Strick, P. L. (1979). Frontal lobe inputs to primate motor cortex: Evidence for four somatotopically organized 'premotor' areas. *Brain Research, 177*(1), 176-182.

Murray, F. P. (1969). An investigation of variably induced white noise upon moments of stuttering. *Journal of Communications Disorders, 2*, 109-114.

Mushiake, H., & Strick, P. L. (1995). Pallidal neuron activity during sequential arm movements. *Journal of Neurophysiology, 74*(6), 2754-2758.

Mysak, E. D. (1960). Servo theory and stuttering. *Journal of Speech and Hearing Disorders, 25*, 188-195.

Neilson, M. D., & Neilson, P. D. (1987). Speech motor control and stuttering: A computational model of adaptive sensory-motor processing. *Speech Communication, 6*(4), 283-373.

Neilson, M. D., & Neilson, P. D. (1991). Adaptive model theory of speech motor control and stuttering. In H. F. M. Peters, W. Hulstijn & C. W. Starkweather (Eds.), *International Conference on Speech Motor Control and Stuttering (2nd : 1990 : Nijmegen, Netherlands). Speech motor control and stuttering* (pp. 149-156). Amsterdam: Excerpta Medica.

Neumann, K., Euler, H. A., von Gudenberg, A. W., Giraud, A. L., Lanfermann, H., Gall, V., et al. (2003). The nature and treatment of stuttering as revealed by fMRI a within- and between-group comparison. *Journal of Fluency Disorders, 28*(4), 381-409.

Nevet, A., Morris, G., Saban, G., Fainstein, N., & Bergman, H. (2004). Discharge rate of substantia nigra pars reticulata neurons is reduced in non-parkinsonian monkeys with apomorphine-induced orofacial dyskinesia. *Journal of Neurophysiology, 92*(4), 1973-1981.

Nieto-Castanon, A., Guenther, F. H., Perkell, J. S., & Curtin, H. D. (2005). A modeling investigation of articulatory variability and acoustic stability during American English /r/ production. *Journal of the Acoustical Society of America, 117*(5), 3196-3212.

Nippold, M. A. (2002). Stuttering and phonology: Is there an interaction? *American Journal of Speech-Language Pathology, 11*, 99–110.

Nudelman, H. B., Herbrich, K. E., Hoyt, B. D., & Rosenfield, D. B. (1987). Dynamic characteristics of vocal frequency tracking in stutterers and nonstutterers. In H. F. M. Peters & W. Hulstijn (Eds.), *Speech Motor Dynamics in Stuttering (1985 : University of Nijmegen)*. *Speech motor dynamics in stuttering* (pp. 161-169). New York: Springer-Verlag.

Parent, A., & Hazrati, L. N. (1995). Functional anatomy of the basal ganglia. II. The place of subthalamic nucleus and external pallidum in basal ganglia circuitry. *Brain Research. Brain Research Reviews, 20*(1), 128-154.

Parent, M., & Parent, A. (2006). Single-axon tracing study of corticostriatal projections arising from primary motor cortex in primates. *Journal of Comparative Neurology, 496*(2), 202-213.

Parker, C. S., & Christopherson, F. (1963). Electronic aid in the treatment of stammer. *Medical Electronics and Biological Engineering, 1*, 121-125.

Passingham, R. (1993). *The frontal lobes and voluntary activity*. Oxford: Oxford U. Press.

Peeva, M. G., Guenther, F. H., Tourville, J. A., Nieto-Castanon, A., Anton, J. L., Nazarian, B., et al. (2009). *Distinct representations of phonemes, syllables, and supra-syllabic sequences in the speech production network*. Manuscript submitted for publication.

Percheron, G., Yelnik, J., & Francois, C. (1984). A golgi analysis of the primate globus pallidus. III. Spatial organization of the striato-pallidal complex. *Journal of Comparative Neurology, 227*(2), 214-227.

Perkell, J. S., Guenther, F. H., Lane, H., Matthies, M. L., Stockmann, E., Tiede, M., et al. (2004). The distinctness of speakers' productions of vowel contrasts is related to their discrimination of the contrasts. *Journal of the Acoustical Society of America, 116*(4 Pt 1), 2338-2344.

Perkell, J. S., Matthies, M. L., Tiede, M., Lane, H., Zandipour, M., Marrone, N., et al. (2004). The distinctness of speakers' /s/-/S/ contrast is related to their auditory discrimination and use of an articulatory saturation effect. *Journal of Speech, Language, and Hearing Research, 47*(6), 1259-1269.

Perkins, W. H., Bell, J., Johnson, L., & Stocks, J. (1979). Phone rate and the effective planning time hypothesis of stuttering. *Journal of Speech and Hearing Research, 22*(4), 747-755.

Peterson., G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *The Journal of the acoustical society of America, 24. N° 2*, 175-184.

Poletto, C. J., Verdun, L. P., Strominger, R., & Ludlow, C. L. (2004). Correspondence between laryngeal vocal fold movement and muscle activity during speech and nonspeech gestures. *Journal of Applied Physiology, 97*(3), 858-866.

Postma, A. (2000). Detection of errors during speech production: A review of speech monitoring models. *Cognition, 77*(2), 97-132.

Postma, A., & Kolk, H. (1992). Error monitoring in people who stutter: Evidence against auditory feedback defect theories. *Journal of Speech and Hearing Research, 35*(5), 1024-1032.

Postma, A., & Kolk, H. (1993). The covert repair hypothesis: Prearticulatory repair processes in normal and stuttered disfluencies. *Journal of Speech and Hearing Research, 36*(3), 472-487.

Prosek, R. A., Montgomery, A. A., Walden, B. E., & Hawkins, D. B. (1987). Formant frequencies of stuttered and fluent vowels. *Journal of Speech and Hearing Research, 30*(3), 301-305.

Purcell, D. W., & Munhall, K. G. (2006). Adaptive control of vowel formant frequency: Evidence from real-time formant manipulation. *Journal of the Acoustical Society of America, 120*(2), 966-977.

Riley, G. D. (1994). *Stuttering severity instrument for children and adults* (3rd ed.). Austin, TX: PRO-ED.

Robb, M., & Blomgren, M. (1997). Analysis of F2 transitions in the speech of stutterers and nonstutterers. *Journal of Fluency Disorders, 22*(1), 1-16.

Russell, M., Corley, M., & Lickley, R. (2005). Magnitude estimation of disfluency by stutterers and nonstutterers. In R. B. R. Hartsuiker, A. Postma, & F. Wijnen (Ed.). *Phonological encoding and monitoring in normal and pathological speech* (pp. 248–260). Hove, UK: Psychology Press.

Saltuklaroglu, T., & Kalinowski, J. (2006). The inhibition of stuttering via the presentation of natural speech and sinusoidal speech analogs. *Neuroscience Letters, 404*(1-2), 196-201.

Scott, C. M., & Ringel, R. L. (1971). Articulation without oral sensory control. *Journal of Speech and Hearing Research, 14*(4), 804-818.

Shane, M. L. S. (1955). Effects on stuttering of alteration in auditory feedback. In W. S. Johnson & R. R. Leutenegger (Eds.). *Stuttering in children and adults: Thirty years of research at the university of Iowa* (pp. 286-297). Minneapolis: University of Minnesota Press.

Sheehan, J. G. (1974). Stuttering behavior: A phonetic analysis. *Journal of Communication Disorders, 7*, 193-212.

Sherrard, C. A. (1975). Stuttering as "false alarm" responding. *British Journal of Disorders of Communication, 10*(2), 83-91.

Smith, A. (1992). The control of orofacial movements in speech. *Critical Reviews in Oral Biology and Medicine, 3*, 233-267.

Smith, A. (1999). Stuttering: A unified approach to a multifactorial, dynamic disorder. In N. B. Ratner & E. C. Healey (Eds.). *Stuttering research and practice: Bridging the gap* (pp. 27-43). Mahwah, NJ: Erlbaum.

Soderberg, G. A. (1967). Linguistic factors in stuttering. *Journal of Speech and Hearing Research, 10*(4), 801-810.

Sommer, M., Koch, M. A., Paulus, W., Weiller, C., & Buchel, C. (2002). Disconnection of speech-relevant brain areas in persistent developmental stuttering. *Lancet, 360*(9330), 380-383.

Stager, S. V., Calis, K., Grothe, D., Bloch, M., Berensen, N. M., Smith, P. J., et al. (2005). Treatment with medications affecting dopaminergic and serotonergic mechanisms: Effects on fluency and anxiety in persons who stutter. *Journal of Fluency Disorders, 30*(4), 319-335.

Stahl, S. M. (2001). Dopamine system stabilizers, aripiprazole, and the next generation of antipsychotics, part 2: Illustrating their mechanism of action. *Journal of Clinical Psychiatry, 62*(12), 923-924.

Starkweather, C. W. (1987). *Fluency and stuttering*. Englewood Cliffs, NJ: Prentice-Hall.

Stephen, S. C. G., & Haggard, M. P. (1980). Acoustic properties of masking/delayed feedback in the fluency of stutterers and controls. *Journal of Speech and Hearing Research, 23*, 527-538.

Stevens, K. N. (1998). *Acoustic phonetics*. Cambridge, MA: The MIT Press.

Stromsta, C. (1959). Experimental blockage of phonation by distorted sidetone. *Journal of Speech and Hearing Research, 2*, 286-301.

Stromsta, C. (1972). Interaural phase disparity of stutterers and nonstutterers. *Journal of Speech and Hearing Research, 15*(4), 771-780.

Stromsta, C. (1986). *Elements of stuttering*. Oshtemo, MI: Atsmorts Publishing.

Stromsta, C., & Fibiger, S. (1981). Physiological correlates of the core behavior of stuttering. In B. J. Urban (Ed.), *18th Congress of the International Association of Logopedics and Phoniatrics : Washington, DC, 1980* (pp. 335-340). Rockville, MD: American Speech-Language-Hearing Association.

Subramanian, A., Yairi, E., & Amir, O. (2003). Second formant transitions in fluent speech of persistent and recovered preschool children who stutter. *Journal of Communications Disorders, 36*, 59-75.

Sutton, S., & Chase, R. A. (1961). White noise and stuttering. *Journal of Speech and Hearing Research, 4*, 72.

Takada, M., Tokuno, H., Nambu, A., & Inase, M. (1998). Corticostriatal projections from the somatic motor areas of the frontal cortex in the macaque monkey: Segregation versus overlap of input zones from the primary motor cortex, the supplementary motor area, and the premotor cortex. *Experimental Brain Research, 120*(1), 114-128.

Tasko, S. M., McClean, M. D., & Runyan, C. M. (2007). Speech motor correlates of treatment-related changes in stuttering severity and speech naturalness. *Journal of Communication Disorders, 40*(1), 42-65.

Tourville, J. A., Reilly, K. J., & Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *Neuroimage, 39*(3), 1429-1443.

Toyomura, A., & Omori, T. (2004). Stochastic control system of basal ganglia and relation with stuttering. In M. H. Hamza (Ed.), *The 2nd IASTED International Conference on Neural Networks and Computational Intelligence : Grindelwald, Switzerland* (pp. 31-36). Anaheim, CA: Acta Press.

Tran, N. L., Maguire, G. A., Franklin, D. L., & Riley, G. D. (2008). Case report of aripiprazole for persistent developmental stuttering. *Journal of Clinical Psychopharmacology, 28*(4), 470-472.

Ueda, Y., & Kimura, M. (2003). Encoding of direction and combination of movements by primate putamen neurons. *European Journal of Neuroscience, 18*(4), 980-994.

van Lieshout, P. H. H. M. (2004). Searching for the weak link in the speech production chain of people who stutter: A motor skill approach. In B. Maassen, R. D. Kent, H. F. M. Peters, P. H. H. M. van Lieshout & W. Hulstijn (Eds.). *Speech motor control in normal and disordered speech*. Oxford, UK: Oxford University Press.

van Lieshout, P. H. H. M., Hulstijn, W., & Peters, H. F. (1996a). From planning to articulation in speech production: What differentiates a person who stutters from a person who does not stutter? *Journal of Speech and Hearing Research, 39*(3), 546-564.

van Lieshout, P. H. H. M., Hulstijn, W., & Peters, H. F. (1996b). Speech production in people who stutter: Testing the motor plan assembly hypothesis. *Journal of Speech and Hearing Research, 39*(1), 76-92.

van Lieshout, P. H. H. M., Peters, H. F. M., Starkweather, C. W., & Hulstijn, W. (1993). Physiological differences between stutterers and nonstutterers in perceptually fluent speech: EMG amplitude and duration. *Journal of Speech and Hearing Research, 36*, 55-63.

Van Riper, C. (1971). *The nature of stuttering*. Englewood Cliffs, NJ: Prentice-Hall.

Van Riper, C. (1982). *The nature of stuttering*. Englewood Cliffs, NJ: Prentice-Hall.

Venkatagiri, H. S. (1980). The core disruption under delayed auditory feedback: Evidence from adaptation study. *Journal of Communication Disorders, 13*(5), 365-371.

Villacorta, V. M., Perkell, J. S., & Guenther, F. H. (2007). Sensorimotor adaptation to feedback perturbations on vowel acoustics and its relation to perception. *Journal of the Acoustical Society of America, 122*(4), 2306-2319.

Vollenweider, F. X., Maguire, R. P., Leenders, K. L., Mathys, K., & Angst, J. (1998). Effects of high amphetamine dose on mood and cerebral glucose metabolism in normal volunteers using positron emission tomography (PET). *Psychiatry Research, 83*(3), 149-162.

Watkins, K. E., Smith, S. M., Davis, S., & Howell, P. (2008). Structural and functional abnormalities of the motor system in developmental stuttering. *Brain, 131*(Pt 1), 50-59.

Webster, R. L., & Dorman, M. F. (1970). Decreases in stuttering frequency as a function of continuous and contingent forms of auditory masking. *Journal of Speech and Hearing Research, 13*(1), 82-86.

Webster, R. L., & Lubker, B. B. (1968). Masking of auditory feedback in stutterers' speech. *Journal of Speech and Hearing Research, 11*(1), 221-223.

Wichmann, T., & Delong, M. R. (1996). Functional and pathophysiological models of the basal ganglia. *Current Opinion in Neurobiology, 6*(6), 751-758.

Wilson, C. J. (1995). The contribution of cortical neurons to the firing pattern of striatal spiny neurons. In J. C. Houk, J. L. Davis & D. G. Beiser (Eds.). *Models of information processing in the basal ganglia*: The MIT Press.

Wilson, C. J. (2004). Basal ganglia. In G. M. Shepherd (Ed.). *The synaptic organization of the brain* (5th ed. ed.): Oxford University Press.

Wingate, M. E. (1964). A standard definition of stuttering. *Journal of Speech and Hearing Disorders, 29*, 484-489.

Wingate, M. E. (1969). Stuttering as phonetic transition defect. *Journal of Speech and Hearing Disorders, 34*, 107-108.

Wingate, M. E. (1976). *Stuttering: Theory and treatment*. New York: Irvington Publishers.

Wingate, M. E. (2002). *Foundations of stuttering*. San Diego, CA: Academic Press.

Wolkin, A., Angrist, B., Wolf, A., Brodie, J., Wolkin, B., Jaeger, J., et al. (1987). Effects of amphetamine on local cerebral metabolism in normal and schizophrenic subjects as determined by positron emission tomography. *Psychopharmacology, 92*(2), 241-246.

Wu, J. C., Maguire, G., Riley, G., Fallon, J., Lacasse, L., Chin, S., et al. (1995). A positron emission tomography [18F]deoxyglucose study of developmental stuttering. *Neuroreport, 6*(3), 501-505.

Wu, J. C., Maguire, G., Riley, G., Lee, A., Keator, D., Tang, C., et al. (1997). Increased dopamine activity associated with stuttering. *Neuroreport, 8*(3), 767-770.

Yairi, E. (1976). Effects of binaural and monaural noise on stuttering. *The journal of Auditory Research, 16*, 114-119.

Yairi, E., & Ambrose, N. G. (2005). *Early childhood stuttering for clinicians by clinicians*. Austin, TX: PRO-ED.

Yaruss, J. S., & Conture, E. G. (1993). F2 transitions during sound/syllable repetitions of children who stutter and predictions of stuttering chronicity. *Journal of Speech and Hearing Research, 36*(5), 883-896.

Yates, A. J. (1963). Delayed auditory feedback. *Psychological Bulletin, 60*, 213-232.

Zikopoulos, B., & Barbas, H. (2007). Parallel driving and modulatory pathways link the prefrontal cortex and thalamus. *PLoS ONE, 2*(9), e848.

Zimmermann, G. (1980a). Articulatory behaviors associated with stuttering: A cinefluorographic analysis. *Journal of Speech and Hearing Research, 23*(1), 108-121.

Zimmermann, G. (1980b). Stuttering: A disorder of movement. *Journal of Speech and Hearing Research, 23*, 122-136.

CURRICULUM VITAE

Name   Oren Civier

Year of Birth   1976

Contact Address   Kehilat Vilna 13

Ramat Hasharon, 47220, ISRAEL

Email   orencivboston@gmail.com

Education

Boston University   *Ph.D., Cognitive and Neural Systems*

Boston, Massachusetts   Jan. 2010

Advisor: Prof. Frank Guenther

Dissertation: *Computational Modeling of the Neural Substrates of Stuttering and Induced Fluency*

(can be found at this URL: http://speechlab.bu.edu/publications.php#dissertations)

Open University of Israel   *B.A. (cum laude), Computer Science*

Tel Aviv, Israel   Nov. 2000

Publications

Civier, O., Bullock, D. H., Max, L., and Guenther, F. H. (in preparation) Basal ganglia dysfunction may lead to dysfluencies: simulating neural impairments to syllable-level command generation in stuttering.

Civier, O., Tasko, S. M., and Guenther, F. H. (accepted) Overreliance on auditory feedback may lead to sound/syllable repetitions: simulations of stuttering and fluency-inducing conditions with a neural model of speech production. *Journal of Fluency Disorders.*

Civier, O., Bullock, D. H., Max, L., and Guenther, F. H. (2009) Simulating neural impairments to syllable-level command generation in stuttering. *Proceedings of the 6th World Congress on Fluency Disorders.* Rio de Janeiro, Brazil, 5th to 8th August, 2009.

Civier, O., Tasko, S. M., and Guenther, F. H. (2009) Overreliance on auditory feedback control in stuttering: A modeling study. *Proceedings of the 6th World Congress on Fluency Disorders.* Rio de Janeiro, Brazil, 5th to 8th August, 2009.

Civier, O., and Guenther, F. H. (2005). Simulations of feedback and feedforward control in stuttering. *Proceedings of the 7th Oxford Dysfluency Conference.* St. Catherine's College, Oxford, 29th June to 2nd July, 2005.

Civier, O., and Guenther, F. H. (2005). Simulations of feedback and feedforward control in stuttering. *Poster presented at the 9th International Conference on Cognitive and Neural Systems.* Boston University, 18th to 21st May, 2005.

Experience

Boston University   *Research Assistant in the Speech Lab,*

                                       Department of Cognitive & Neural Systems (CNS)

Boston, Massachusetts   Sep. 2008 – Mar. 2009

Boston University   *Research assistant in the Speech Lab, and*

                                       Teaching Assistant in the CNS department

Boston, Massachusetts   June 2003 – Dec. 2006

Israeli Defense Force   *Information Technology Project Manager*

Israel   Nov. 1998 – May 2001

Israeli Defense Force   *System Administrator and Programmer*

Israel   July 1995 – Oct. 1998