Interspeech 2025
17-21 August 2025, Rotterdam, The Netherlands

Towards a dynamical model of transitions between fluent and stuttered speech

Yijing Lu[1*], *Khalil Iskarous*[2], *Louis Goldstein*[2]

[1]University of Potsdam, Germany
[2]University of Southern California, United States

yijing.lu@uni-potsdam.de, kiskarou@usc.edu, louisgol@usc.edu

Abstract

This paper introduces a dynamical systems framework for understanding stuttering, conceptualizing it as a qualitative shift in speech articulation driven by a single control parameter. Using a forced Duffing oscillator model, we demonstrate how variations in the excitation frequency can account for transitions between fluent and stuttered speech states. The model generates specific predictions about articulatory behaviors during stuttering, which we test using real-time MRI data of stuttered speech. Analysis of articulatory movements provides empirical support for the model's predictions, suggesting that stuttering can be understood as a dynamical disease—an intact system operating outside its typical parameter range. This framework offers new insights into the nature of stuttering and potential approaches to intervention.

Index Terms: stuttering, articulation, dynamical disease, dynamical system theory, hysteresis, Duffing oscillator

1. Introduction

Repetitive articulatory movements are a defining characteristic of stuttered speech [1, 2, 3] and represent the earliest behavioral marker in the developmental trajectory of stuttering [4]. Thus, it is not an overstatement to say they are the core behavioral feature of developmental stuttering, constituting a key part of its definition. However, possibly due to the historical lack of articulatory data on stuttered speech, these repetitive movements remain insufficiently studied, let alone modeled. In this paper, we apply dynamical systems theory to model the emergence and dissolution of these repetitive movements, using a forced oscillator framework whose predictions we test with articulatory data.

1.1. Dynamical systems theory approach to stuttering

Dynamical systems theory is a mathematical framework that uses differential equations to model phenomena that evolve over time and/or space, providing insights into the nature of these changes. It has been applied in various fields such as physics, biology, economics, engineering, ecology, and also speech motor control [5]. In 1977, Mackey and Glass introduced this approach to the study of human pathologies and coined the term "dynamical disease" [6]. A dynamical disease involves changes in the qualitative dynamics of some observable processes as one or more control parameters are changed. Pathological states are thus viewed as the result of a normal physiological control system operating outside their typical range of parameters, suggesting the possibility of returning to a healthy state by adjusting these parameters. Over the past several decades, the framework of dynamical diseases has been applied to various pathological conditions, such as cardiac and respiratory arrhythmias, epileptic seizures, periodic hematological diseases, bipolar disorder, and schizophrenia, yielding valuable insights into their nature, diagnostics and treatment [7, 8].

Stuttering is well-suited to be studied under the framework of dynamical diseases due to several of its fundamental characteristics. [9] identified three types of qualitative changes in dynamics associated with dynamical diseases: "(1) the appearance of a regular oscillation in a physiological control system not normally characterized by rhythmic processes, (2) the development of new periodicity in an already periodic process, and (3) the disappearance of a rhythmic process" (p. 17). Stuttering aligns with the first category, as the transition from fluent speech to stuttering involves a qualitative shift in speech production: in fluent speech, articulatory movements are typically viewed as a sequence of goal-directed, discrete actions [10]; during moments of stuttering, these movements become repetitive and can instead be viewed holistically as oscillation.

Importantly, the essence of a dynamical disease is an intact physiological control system operating within an abnormal range of control parameters. In the case of stuttering, there is no identified difference in speech musculature between people who stutter (PWS) and those who do not (PWNS) [11]. While some structural and functional brain differences have been observed between PWS and PWNS [12], these differences do not necessarily indicate a pathological speech motor control system. Instead, they could reflect compensatory adaptations or developmental variations that allow PWS to manage the demands of speech production. [13] argued that PWS occupy the lower end of a speech motor skill continuum rather than having a qualitative difference from PWNS. It is also important to keep in mind that stuttering is an intermittent phenomenon, as PWS are capable of producing perceptually fluent speech. Taken all together, these factors suggest that the speech motor control system of PWS is possibly intact but operates within an abnormal range of control parameters during stuttering episodes. Identifying the critical parameter(s) responsible for triggering stuttering could open up new possibilities for treatment.

Although it has been suggested that dynamical systems theory offers a framework for understanding stuttering [14, 15, 16], very few studies have explored this approach in depth. Existing research includes modeling work on trait differences in speech rhythm between PWS and PWNS [17, 18, 19] and empirical investigations examining coordination dynamics in fluent speech of both groups [20, 21]. No studies have applied dynamical systems theory to model the transition between fluent and disfluent states in the speech of PWS. This paper begins this important effort by proposing an explicit model and testing its predictions.

[*]Work primarily conducted during PhD at the University of Southern California.

10.21437/Interspeech.2025-2320

1.2. Nonlinear modeling of speech movement

The primary mathematical tools for modeling dynamical diseases are nonlinear dynamical systems, as nonlinear systems can exhibit qualitative changes in behavior. In the context of typical speech, the classic model for articulatory movement employs a linear system [22] (Equation 1). However, because this linear system fails to accurately predict the velocity profiles and kinematic relationships observed in articulatory experiments, recent work [23] has proposed adding an nonlinear term to improve the model's fit to real kinematic data (Equation 2).

$$\ddot{x} + b\dot{x} + kx = 0 \quad (b > 0, k > 0) \tag{1}$$

$$\ddot{x} + b\dot{x} + kx - dx^3 = 0 \quad (b > 0, k > 0, d > 0) \tag{2}$$

In fact, Equation 2 is a special case of the Duffing oscillator [24], a well-known nonlinear dynamical system. A standard Duffing oscillator is driven by an external periodic force and can be expressed as:

$$\ddot{x} + \delta\dot{x} + \alpha x + \beta x^3 = \gamma \cos(\omega t) \tag{3}$$

Here, x is the displacement, δ is the damping coefficient, α is the linear stiffness coefficient, β is the nonlinear stiffness coefficient, γ is the amplitude of the periodic forcing, and ω is the angular frequency of the periodic forcing. While [23] set the external forcing term γ to 0, we propose reintroducing a nonzero forcing term ($\gamma > 0$) to model both fluent and stuttered speech, as we will elaborate in Section 1.3.

This modification addresses key limitations of the unforced Duffing oscillator in modeling stuttering. Without periodic forcing, the system can only produce oscillations under zero damping, a condition that is unrealistic in biological systems. Furthermore, the amplitude of the oscillation generated in such a scenario depend on the initial state, which does not align with observed characteristics of stuttering (e.g., the overall amplitude of lip oscillation during stuttering does not appear to be determined by the degree of lip closing in the initial half-cycle). A forced Duffing oscillator, on the other hand, may offer a more promising model for capturing the complex oscillatory behavior seen in stuttering, where the force act as an external trigger for stuttering (the interpretation of the external force will be discussed in Section 5).

1.3. Forced Duffing oscillator

In this paper, we propose using a forced Duffing oscillator to model articulatory oscillation during stuttering. The same model is also hypothesized to underlie fluent speech, as we will demonstrate shortly. Like [23], we assume a monostable, softening Duffing oscillator (i.e., $\alpha > 0$, $\beta < 0$). As a nonlinear oscillator, it exhibits distinct behaviors compared to linear systems in response to external forces, allowing us to generate predictions for articulatory behaviors during stuttering.

First, resonance of nonlinear systems does not occur when the frequency of the external force (i.e., excitation frequency, ω) equals the natural frequency of the system (ω_0). As shown in Figure 1a, the frequency response function (i.e., amplitude of the steady-state oscillation as a function of the excitation frequency) of a forced Duffing oscillator has its peak tilted towards the lower-frequency side when $\beta < 0$ and towards the higher-frequency side when $\beta > 0$. Regardless of the value of β, the frequency of the response oscillation generally equals the excitation frequency, as shown in Figure 1b.

Additionally, in forced Duffing oscillators, the amplitude of response oscillation exhibits discontinuity when the excitation

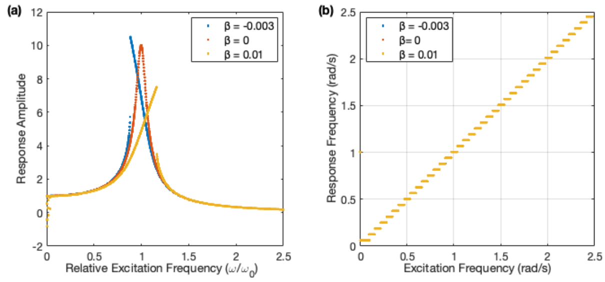

Figure 1: *(a) Amplitude and (b) frequency of steady-state oscillation as a function of excitation frequency. Parameters:* $\alpha = 1$, $\gamma = 1$, $\delta = 0.1$, $\beta = [-0.003, 0, 0.01]$.

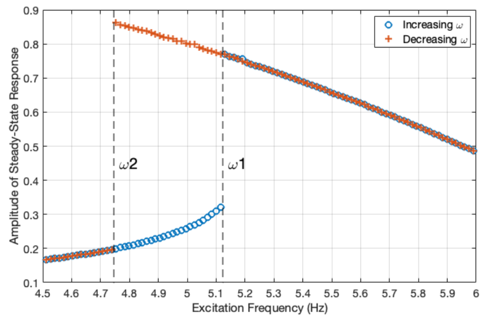

Figure 2: *Simulated frequency response function for increasing and decreasing ω in the range of 4.5 to 6 Hz. Parameters: see Section 4.*

frequency (ω) crosses a critical value, as shown by the sudden jumps in the response functions in Figure 1a. For $\beta < 0$ (the condition we are interested in), the amplitude begins to increase as ω rises, then suddenly jumps up when ω reaches a critical value. After that point, the amplitude decreases as ω continues to increase. If we assume that the lower amplitude oscillation before the critical point of ω exists in fluent speech but is not observable in articulation, the sudden jump up could be seen as the moment stuttering occurs—the oscillation becomes significant and observable in articulation. Thus, we hypothesize that the excitation frequency (ω) is the control parameter for triggering stuttering.

Another important phenomenon in forced Duffing oscillators is hysteresis, which means the change of the system not only depends on the current state but also on its history. As shown in Figure 2, when ω slowly increases from the lower end, the response amplitude jumps up at a critical frequency, ω_1. However, when ω slowly decreases from the higher end, the sudden jump down in response amplitude occurs at a different frequency, ω_2. The critical value of ω depends on the direction of frequency change.

Based on these characteristics, we can form several predictions for articulatory oscillations during stuttering. We suggest that entering a stuttering state corresponds to an increase in ω beyond a critical value, ω_1, and exiting the stuttering state requires a decrease in ω. Since the frequency of the resulting oscillation is the same as the excitation frequency (ω), the **first prediction** is that the frequency of articulatory oscillation decreases towards the end of stutters. Then, given the shape of the frequency response function, as ω decreases, the amplitude of the response oscillation actually increases. So, the **second prediction** is that the amplitude of articulatory oscillation increases towards the end of stutters. The stuttering state will persist until

ω reaches ω_2, the point at which the response amplitude suddenly drops. Note that the magnitude of the jump up at ω_1 is smaller than the magnitude of the jump down at ω_2. Therefore, the **third prediction** is that the change in the oscillation amplitude when entering a stutter is smaller than the change when exiting.

We test these predictions in the articulatory data of one adult who stutters, with methods described in Section 2. The results are presented in Section 3. A toy model of stuttering using forced Duffing oscillator is presented in Section 4. Section 5 discusses the results and the model.

2. Methods

The participant is a 47-year-old adult who stutters. The articulatory data are a subset of real-time MRI speech production videos of adults who stutter [25], captured using a 0.55T MRI system (Siemens Aera XQ) in the mid-sagittal plane. The reconstructed videos have a spatial resolution of 2.3 × 2.3 mm^2 and a temporal resolution of 10.06 ms per frame (99.4 frames per second). For this study, we focus on the lip oscillations during perceived repetitions of /b, p, m/ and tongue tip oscillations during perceived repetitions of /t, d, n/. Two circular regions-of-interest (radius: 4 pixels) were placed respectively between the lips and under the alveolar ridge to capture pixel intensity changes, which serve as proxies for changes in constriction degree at these two regions. The yielded pixel intensity time series were smoothed using a wavelet transform-based method to remove small fluctuations caused by MRI image noise instead of articulator movement. Time series were then trimmed based on 20% velocity threshold, yielding the intervals from the onset of the closing movement, through the oscillatory part, to the offset of the opening movement that connects to the following fluent speech (Figure 3a).

Since the model and predictions need to be tested in signals with clear oscillatory characteristics, we first selected tokens based on their periodicity metrics. Specifically, we computed the autocorrelation of each signal. Only tokens that exhibit at least two peaks in the autocorrelation function were retained, resulting in a total of 83 valid tokens (66 lip oscillation, 17 tongue tip oscillation). Then we performed spectral analysis to determine the dominant frequency, focusing on the 1–10 Hz range for interpretability, as these signals reflect articulator movements. Each signal was mean-centered to remove the DC offset and tapered with a Hanning window of the full signal length. A 512-point Fast Fourier Transform was applied. The dominant frequency was identified as the highest peak in the spectrum within the 1–10 Hz range. Then, since signals do not always oscillate around a steady mean, we computed the moving mean using the one cycle of the dominant frequency as the window size to detrend the signal. Over the detrended signals, we then applied Hilbert transform to extract instantaneous frequencies and their corresponding amplitudes. To test predictions 1 and 2, we analyzed the temporal evolution of articulatory oscillations by fitting linear regression lines to the instantaneous frequencies and amplitudes within the final 14% of each signal (Figure 3c, 3d). This window was selected based on systematic exploration of different temporal windows (ranging from 10% to 50% of signal duration), which revealed that this portion optimized the linear fits for both frequency and amplitude trends, as measured by the product of the median adjusted R^2 values for amplitude and frequency fit across all tokens (see Section 3.2). To test prediction 3, we compared the amplitude of movement when entering and exiting stuttering (shaded areas in Figure 3a). As a baseline, we also compared the amplitude of the closing movement and opening movement in the participant's fluent articulation of same phonemes.

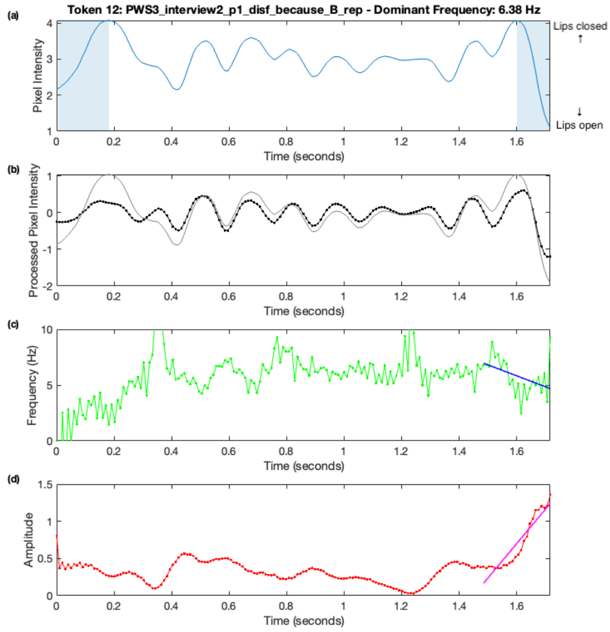

Figure 3: *(a) Pixel intensity signal presenting lip opening and closing over time. (b) Processed signal after mean-centering (gray) and detrending (black). (c) Instantaneous frequencies of the detrended signal and (d) corresponding amplitudes.*

3. Results

3.1. Frequency of articulator oscillation during stuttering

With lip oscillation and tongue tip oscillation tokens pooled together, the dominant frequency falls in the range of 5.5-6 Hz. (Figure 4). This finding is consistent with [26], who found peak frequencies of 5-6 Hz in lower lip oscillations during bilabial disfluencies.

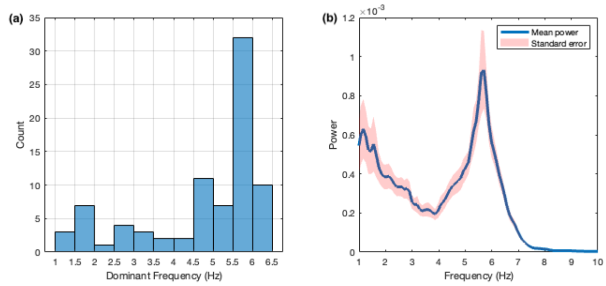

Figure 4: *(a) Histogram of dominant frequencies. (b) Average power spectrum.*

3.2. Change of frequency and amplitude during stuttering

Within the final 14% of the signals, 60% of the examined tokens exhibit both a decreasing trend in instantaneous frequencies, and an increasing trend in the amplitudes. The portion of the signal was chosen because it yielded the highest combined model fits for frequency and amplitude, based on the product of median adjusted R^2 values (Figure 5).

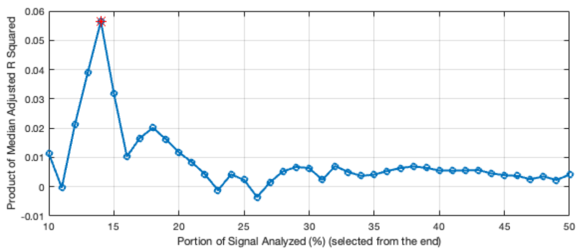

Figure 5: *Relationship between portion of signal analyzed and the product of median adjusted R^2 values.*

3.3. Movement amplitudes entering and exiting stuttering

The amplitude of lip movement out of oscillation was found to be significantly greater than the amplitude of lip movement into oscillation (66 paired observations, Wilcoxon signed rank test: $p < 0.001, W = 344$), and the same was observed for tongue tip oscillation (17 paired observations, Wilcoxon signed rank test: $p < 0.001, W = 0$) (Figure 6). For fluent speech, no significant difference was observed in the movement amplitude into and out of a constriction target for lip gestures (48 paired observations; Wilcoxon signed rank test: $p = 0.07, W = 410$) and tongue tip gestures (8 paired observations; Wilcoxon signed rank test: $p = 0.74, W = 15$).

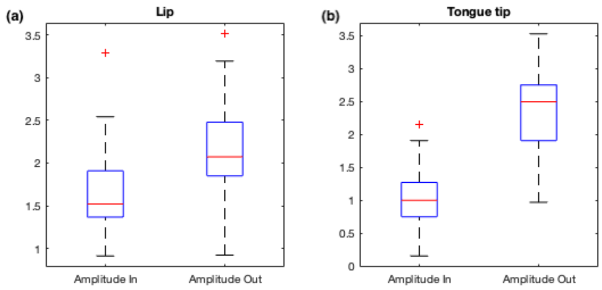

Figure 6: *(a) Lip movement amplitudes and (b) tongue tip movement amplitudes into and out of oscillation.*

4. Toy model

Given that stuttering oscillation exhibits its strongest dominant frequency at approximately 5.5 Hz, it is desirable to have large-amplitude oscillations around 5.5 Hz in the simulation. Since resonance happens at a frequency lower than the natural frequency (see Figure 1), we estimate the natural frequency of the Duffing oscillator as 6 Hz. With the natural frequency established, we can determine the value of α, which is $(2\pi \cdot 6)^2 = 1412.2$. To create a suitable frequency response function, we then set $\beta = -1100$, $\delta = 2$, $\gamma = 100$. The resulting frequency response function in the range of 4.5 to 6 Hz is shown in Figure 2. When ω is increasing, the critical value where the amplitude of response oscillation jumps up is around 5.12 Hz (ω_1), whereas when ω is decreasing, the amplitude jumps down around 4.74 Hz (ω_2).

The transition from fluent speech to stuttered speech and then back to fluent speech is hypothesized as a result of first increasing ω and then decreasing ω. A simulated time series illustrating the transition dynamics is shown in Figure 7. While the simulation uses an extended temporal window to ensure stable system behavior, it demonstrates the key qualitative features of state transitions similar to those seen in stuttering. The amplitude of oscillation significantly increases at the 8th second, marking the start of the stuttering state. This large-amplitude oscillation continues until the 32nd second, indicating the end of the stuttering state. Of particular interest are the two shaded areas (i.e., from the 4th second to the 8th second, and from the 28th second to the 32nd second), where the values of ω are the same but the amplitudes of oscillation differ remarkably due to hysteresis. This hysteresis effect determines that decreasing ω below ω_1 does not immediately terminate the large-amplitude oscillations; instead, the system must pass through a period of increasing amplitude before the sudden drop (shaded area 2).

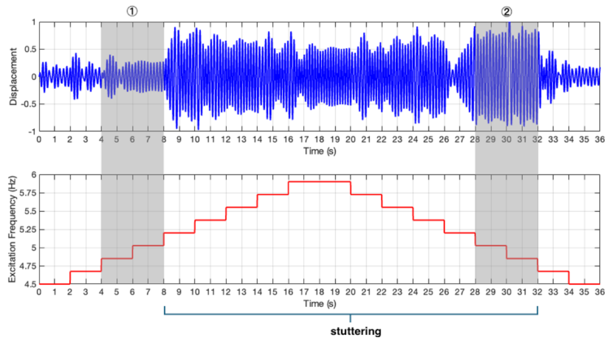

Figure 7: *Duffing oscillator response to time-varying excitation frequency.*

5. Discussion

This paper presents a novel conceptual framework for understanding the transitions between fluent and stuttered speech through a dynamical system perspective. Our forced Duffing oscillator model captures several key qualitative features observed in real articulatory data of stuttering: the sudden onset of observable oscillations, the concurrent changes in oscillation frequency and amplitude during stuttering, and asymmetric transitions into and out of stuttering states.

As the model is still under development, several important questions remain open. First, what does the excitation force correspond to, and why would the excitation frequency remain stably on the lower side for PWNS but easily cross the critical value for PWS? One hypothesis is that the excitation force corresponds to an internal timekeeper for speech. Evidence from both behavioral studies [27, 28, 29] and brain imaging studies [30, 31] suggests that PWS have less stable internal time-keeping. This instability might explain why PWS experience a wider range of excitation frequencies. Second, the relationship between our model and real stuttering episodes requires more consideration. Any forced system undergoes an initial transient phase before reaching steady-state oscillation, during which the system's behaviors are unsteady and irregular. Some stuttering episodes may contain a larger proportion of transient responses rather than predominantly steady-state behavior, which could explain why not all stuttering tokens show the predicted trends of decreasing frequency and increasing amplitude. Our current simulation uses an extended temporal window to ensure stable system behavior, but real stuttering episodes typically occur on shorter timescales where transient effects may be more prominent. Third, our current simulation focuses on demonstrating fundamental principles of the model rather than providing detailed articulatory trajectories. To achieve a simulation of actual articulatory trajectories, a composite model that integrates discrete movements and oscillation might be necessary. Together, these questions highlight important avenues for future research as we refine our model to better capture the complexities of stuttering.

6. Acknowledgments

We acknowledge support from NSF grants RI-2106930, 182873, 2240349. We are also grateful to Stephan Kuberski and Adamantios Gafos for helpful discussions.

7. References

[1] I. Didirkova, S. Le Maguer, F. Hirsch, and D. Gbedahou, "Articulatory behaviour during disfluencies in stuttered speech," in *ICPhS 2019-19th International Congress on Phonetic Sciences*, 2019, pp. 2991–2995.

[2] Y. Lu, C. E. Wiltshire, K. E. Watkins, M. Chiew, and L. Goldstein, "Characteristics of articulatory gestures in stuttered speech: A case study using real-time magnetic resonance imaging," *Journal of Communication Disorders*, vol. 97, p. 106213, 2022.

[3] K. Teesson, A. Packman, and M. Onslow, "The lidcombe behavioral data language of stuttering," *Journal of Speech, Language & Hearing Research*, vol. 46, no. 4, pp. 1009–1015, 2003.

[4] C. Van Riper, *The Nature of Stuttering*, 2nd ed. Prentice-Hall.

[5] P. H. H. M. Van Lieshout, "Dynamical systems theory and its application in speech," in *Speech Motor Control In Normal and Disordered Speech*, B. Maassen, R. Kent, H. F. M. Peters, P. H. H. M. V. Lieshout, and W. Hulstijn, Eds. Oxford University Press, pp. 51–81.

[6] M. C. Mackey and L. Glass, "Oscillation and chaos in physiological control systems," *Science*, vol. 197, no. 4300, pp. 287–289, 1977.

[7] J. Bélair, F. Nekka, and J. G. Milton, "Introduction to focus issue: Dynamical disease: A translational approach," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 31, no. 6, 2021.

[8] L. Glass, "Dynamical disease: Challenges for nonlinear dynamics and medicine," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 25, no. 9, 2015.

[9] M. C. Mackey and J. G. Milton, "Dynamical diseases," *Annals of the New York Academy of Sciences*, vol. 504, no. 1, pp. 16–32, 1987.

[10] J. A. Kelso, E. L. Saltzman, and B. Tuller, "The dynamical perspective on speech production: Data and theory," *Journal of Phonetics*, vol. 14, no. 1, pp. 29–59, 1986.

[11] F. H. Guenther, *Neural Control of Speech*. MIT Press, 2016.

[12] N. E. Neef and S.-E. Chang, "Knowns and unknowns about the neurobiology of stuttering," *Plos Biology*, vol. 22, no. 2, p. e3002492, 2024.

[13] A. Namasivayam and P. van Lieshout, "Speech motor skill and stuttering." *Journal of Motor Behavior*, vol. 43, no. 6, 2011.

[14] P. J. Alfonso, "Implications of the concepts underlying task-dynamic modeling on kinematic studies of stuttering," in *Speech Motor Control and Stuttering*, H. F. Peters, W. Hulstijn, and C. Starkweather, Eds. Amsterdam: Elsevier, pp. 79–100.

[15] E. Saltzman, "The task dynamic model in speech production," in *Speech Motor Control and Stuttering*, H. F. M. Peters, W. Hulstijn, and C. Starkweather, Eds. Amsterdam: Elsevier, pp. 37–52.

[16] A. Smith and C. Weber, "How stuttering develops: The multifactorial dynamic pathways theory," *Journal of Speech, Language, and Hearing Research*, vol. 60, no. 9, pp. 2483–2505, 2017.

[17] O. Skljarov, "The bifurcation model of the speech rhythm and stuttering," in *Proc. SPM 1996*, 1996, pp. 89–92.

[18] ——, "Nonlinear neurodynamics in representation of a rhythm of speech," *Journal of Biological Physics*, vol. 25, no. 2, pp. 223–234, 1999.

[19] ——, "Neurodynamical route to chaos and normal speech vs. stuttering," in *Proceedings of International Conference on Control of Oscillations and Chaos*, 2000, pp. 449–452.

[20] A. K. Namasivayam and P. van Lieshout, "Investigating speech motor practice and learning in people who stutter," *Journal of Fluency Disorders*, vol. 33, pp. 32–51, 2008.

[21] A. K. Namasivayam, P. Van Lieshout, W. E. McIlroy, and L. De Nil, "Sensory feedback dependence hypothesis in persons who stutter." *Human Movement Science*, vol. 28, no. 6, pp. 688–707, 2009.

[22] E. L. Saltzman and K. G. Munhall, "A dynamical approach to gestural patterning in speech production," *Ecological Psychology*, vol. 1, no. 4, pp. 333–382, 1989.

[23] T. Sorensen and A. Gafos, "The gesture as an autonomous nonlinear dynamical system," *Ecological Psychology*, vol. 28, no. 4, pp. 188–215, 2016.

[24] G. Duffing, *Erzwungene Schwingungen bei veränderlicher Eigenfrequenz und ihre technische Bedeutung*. Vieweg, 1918, no. 41-42.

[25] Y. Lu, D. Byrd, L. Goldstein, and S. Narayanan, "Effective elicitation of stuttering in mri data collection using a suite of connected speech tasks," *Journal of Speech, Language, and Hearing Research*, forthcoming, accepted for publication.

[26] M. McClean, H. Goldsmith, and A. Cerf, "Lower-lip emg and displacement during bilabial disfluencies in adult stutterers," *Journal of Speech, Language, and Hearing Research*, vol. 27, no. 3, pp. 342–349, 1984.

[27] F. R. Boutsen, G. J. Brutten, and C. R. Watts, "Timing and intensity variability in the metronomic speech of stuttering and nonstuttering speakers," *Journal of Speech, Language, and Hearing Research*, vol. 43, no. 2, pp. 513–520, 2000.

[28] S. Falk, T. Müller, and S. Dalla Bella, "Non-verbal sensorimotor timing deficits in children and adolescents who stutter," *Frontiers in Psychology*, vol. 6, p. 847, 2015.

[29] A. G. Sares, M. L. Deroche, D. M. Shiller, and V. L. Gracco, "Timing variability of sensorimotor integration during vocalization in individuals who stutter," *Scientific Reports*, vol. 8, no. 1, p. 16340, 2018.

[30] P. A. Alm, "Stuttering and the basal ganglia circuits: a critical review of possible relations," *Journal of Communication Disorders*, vol. 37, no. 4, pp. 325–369, 2004.

[31] A. C. Etchell, M. Ryan, E. Martin, B. W. Johnson, and P. F. Sowman, "Abnormal time course of low beta modulation in non-fluent preschool children: a magnetoencephalographic study of rhythm tracking," *NeuroImage*, vol. 125, pp. 953–963, 2016.