# VOCAL EFFORT AND WITHIN-SPEAKER COORDINATION IN SPEECH PRODUCTION: EFFECTS ON POSTURAL CONTROL

by

Robert Fuhrman

B.A., University of Oregon, 2010

# A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

## MASTER OF ARTS

in

## THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Linguistics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

December 2014

© Robert Fuhrman, 2014

## Abstract

This thesis probes the joint role of respiration in speech motor control and postural control by examining the effect of increasing the loudness of speech production, or *vocal effort*, on within-speaker coordination. Specifically, this work tested the dual hypothesis that the functional demands of speech production at increasingly higher levels of vocal effort would result in increasingly rigid coordination across multiple bodily subsystems, and that this entrainment would ultimately affect postural control, resulting in a loss of balance

An interactive spontaneous speech task was used to elicit speech at multiple levels of vocal effort by increasing the intended communication distance. Data from acoustic and kinematic measurement domains, including speech, rigid body motion of the head, 2d motion of the body extracted from video, and postural forces and torques measured at the feet, were collected simultaneously. These data were analyzed using a unique collection of techniques for the analysis of non-stationary time-series, which included methods for assessing cross-domain correspondence, system dimensionality, and fluctuation characteristics. The results of these analyses show convergent evidence for both hypotheses. Coordination among kinematic and acoustic measurement domains both strengthens and simplifies at high levels of vocal effort, and evidence of postural instability was found at the highest levels of vocal effort. Subsystem fluctuation characteristics show a direct relationship to the observed effects on coordination, both in terms of their intrinsic properties and in terms of changes due to increased vocal effort.

Although this study did not include a direct measure of respiration, these results highlight the necessity of expanding our understanding of respiration's role in speech motor control, especially insofar as the inevitable crossover between speech and other task domains, such as postural control, is concerned. The methodology used in this study can be straightforwardly expanded towards these ends, and provides a potentially useful in-roads to research in this direction. Even in the absence of a respiratory measure,

these results will be of potential interest to clinicians working on the treatment of patients with speech disorders associated with neurological dysfunction, as occur, for example, in Parkinson's disease.

# Preface

The research contained herein was undertaken as an expansion and reworking of a pilot study published as a conference proceedings by members of the Communication Dynamics Laboratory at UBC prior to my arrival, in Vatikiotis-Bateson, E., Oberg, M., Barbosa, A. V., McAllister, K., Hermiston, N., & Kurth, R. (2009). Postural entrainment by vocal effort in singing and speech. *Proceedings of ESCOM 2009*. Jyväskylä, Finland.

As such, the hypotheses tested in the present study were originally proposed by my supervisor, Eric Vatikiots-Bateson. Also, two of the tools used in this study, *optical flow analysis* and correlation *map analysis*, were developed by a former member of the laboratory, Adriano Vilela-Barbosa. Upon my arrival at the laboratory, I re-analyzed the speech data from the original pilot study using these tools, and designed and carried out the present experiment based on the results of my analysis. The data for the current study were collected primarily by myself, but with some help from other members of the laboratory, who aided in the juggling act of collecting simultaneous motion data using multiple measurement systems. I performed all of the data processing and analysis presented in this thesis.

This research was conducted with the approval of the UBC Behavioral Research Ethics Board, under project number H04-80558: Determining Event Structure in Communicative Interaction.

# **Table of Contents**

Abstract	ii
Preface	iv
Table of Contents	V
List of Tables	ix
List of Figures	X
List of Abbreviations	xii
Acknowledgements	xiii
Dedication	xiv
Chapter 1: Introduction	1
1.1 Vocal Effort	1
1.2 Synchronization in Biological Systems	
1.3 Postural Instability	5
1.4 Non-stationary Time Series and Fluctuations in the Analysis of Co	oordination in
Biological Systems	б
1.5 Summary	
Chapter 2: Methods	9
2.1 Participants	9
2.2 Speaking Task and Procedures	9
2.3 Data Collection, Processing, and Reduction	
2.3.1 Experimental Setup	
2.3.2 Data Collection	
2.3.2.1 Speech	

2.3.2.1	Head Motion	12
2.3.2.3	Torso Motion	13
2.3.2.4	Force Plates	13
2.3.3 Data	a Processing	14
2.3.4 Data	a Reduction	15
2.3.4.1	Speech—F0, acoustic RMS	15
2.3.4.2	Rigid Body Head Motion	15
2.3.4.3	Torso Motion	16
2.3.4.4	Force plates —Center of Pressure (COP)	16
Chapter 3: Ana	lysis & Results	.17
3.1 Analys	sis Techniques	17
3.1.1 Cor	relation Map Analysis (CMA)	17
3.1.2 Sing	gular Value Decomposition (SVD)	20
3.1.3 Detr	rended Fluctuation Analysis (DFA)	23
3.1.4 Prin	cipal Component Analysis (PCA)	31
3.1.4 Stat	istical Analysis	37
3.2 Effects	s of Vocal Effort on Acoustic Signal Magnitudes	37
3.3 Vocal	Effort and Breath Group Period Length	40
3.4 Collec	tive Behavior of the Head, Torso, and Speech Acoustics	42
3.4.1 Fluc	ctuation Analysis of Speech and Affiliated Motion Signals	42
3.4.1.1	General Characteristics of the Scaling Exponent in	
	Each Measurement Domain	43
3.4.1.2	Trends and Differences Across Effort Conditions	44

3.4	.2 Mutual O	Organization of Fluctuations: CMA	46
2	2.3.4.1 FO-2	RMS	49
3	3.4.2.2 Spe	ech Acoustics-Head	49
3	3.4.2.3 Spe	ech Acoustics-Torso	49
	3.4.2.4 Tor	so-Head	50
	3.4.2.5 Ger	neral Remarks	51
3.5	Coordinatio	on with Global COP	51
3.6	Postural Ins	stability	56
3.6	.1 Principal	l component analysis of 6d force plate measures	57
3.6	.2 DFA Res	sults, Individual Forces and Torques	60
Chapte	r 4: Discussio	on	.63
4.1	Increased C	Coordination in 'Shout' Relative to the Lowest Three Effort Conditions	65
4.2	Difference	between 'shout' and 'yell': respiratory perturbation and fluctuation	
	characterist	tics of the Head and Torso	67
4.3	Steps toward	rd a systems view of postural instability induced by increased	
	vocal effort	t	70
Chapte	r 5: Conclusi	ion	.71
5.1	Future Dire	ections	72
Referen	ices		.76
Append	lices		.83
Apper	ndix A: Com	plete statistical results: CMA and SVD	83
A.1	CMA an	d SVD within-subjects ANOVA results	83
A.2	2 CMA an	d SVD HSD results	85

Appendi	x B: Complete statistical results: DFA	. 91
<b>B</b> .1	DFA within-subjects ANOVA results	. 91
B.2	DFA HSD results	. 92
Appendi	x C: Complete statistical results: PCA	. 96
C.1	PCA within-subjects ANOVA results	. 96
C.2	PCA Cumulative variance plots: Individual participants	. 97
C.3	PCA HSD results	. 99

# List of Tables

Fable 2.1 List of 60 Hz time series obtained after post processing and data reduction	
Table 3.1 PCA example one	34
Table 3.2 PCA example two	36
Table 3.3 Force plate PCA results	58

# List of Figures

Figure 2.1 OPTOTRAK headgear	12
Figure 2.2 Example optical flow region of interest.	13
Figure 2.3 Bertec force plates with orientation axes	14
Figure 3.1 Example correlation map	18
Figure 3.2 Correlation map and cross-correlograms	19
Figure 3.3 Correlation maps for idealized signals	22
Figure 3.4 SVD of idealized correlation maps	23
Figure 3.5 DFA of example noise signals	26
Figure 3.6 DFA of signal with power law fluctuations.	28
Figure 3.7 Statistical self-similarity of signal with power law fluctuations	30
Figure 3.8 PCA example one: Three 2 Hz sinusoids	33
Figure 3.9 PCA example two: Three 2 Hz sinusoids, two with added noise	35
Figure 3.10 PCA example two: Original variables and principal components	36
Figure 3.11 Average magnitude of acoustic RMS for individual participants.	38
Figure 3.12 Average Magnitude of F0 (Hz), individual participants	39
Figure 3.13 Average breath cycle period, participant one	41
Figure 3.14 DFA results, head, torso, and speech acoustics	44
Figure 3.15 Power law structure and self-similarity of a torso motion signal	45
Figure 3.16 Acoustic-Kinematic CMA and SVD results	48
Figure 3.17 DFA Results, COP time series	52
Figure 3.18 Center of pressure CMA and SVD results	55
Figure 3.19 CMA and SVD results, COP1-COP2	57

Figure 3.20 PCA Cumulative variance, force plates one and two	59
Figure 3.21 Average squared loadings of force plate variables	60
Figure 3.22 DFA results, force plate component forces and torques	61

# List of Abbreviations

CMA:	Correlation map analysis
SVD:	Singular value decomposition
DFA:	Detrended fluctuation analysis
PCA:	Principal component analysis
OFA:	Optical flow analysis
ANOVA:	Analysis of variance
HSD:	Tukey's honest significant difference
FPL:	Force plate
Fx:	Force measured in the x direction
Fy:	Force measured in the y direction
Fz:	Force measured in the z direction
Mx:	Torque measured in the x direction
My:	Torque measured in the y direction
Mz:	Torque measured in the z direction

# Acknowledgements

Sometimes, it's best to just begin at the beginning. So first, I should acknowledge my Dad, who gave me the final bit of encouragement to accept the offer to come to UBC when I was wavering. That encouragement had innumerable consequences for the path my life has taken since then, this thesis being one of the many goods the decision ultimately brought. Along these lines, thanks are also owed to the rest of my family, who have never ceased to support and encourage me.

At UBC, certainly thanks are owed to too many to count. Starting graduate school with a fantastic cohort of fellow students certainly helped, I couldn't have imagined a better group. And the value of friends like Rob Rivers, Will Carlquist, and Dave Prest can't be overstated.

This thesis never would've happened without the help of Adriano Vilela-Barbosa, Jesse Lawrence, and Graham Haber, for various reasons. I can't thank the three of you enough.

As for the members of my committee, Molly Babel was especially accommodating in helping me to find my way in a new field in the early goings on. I'm sure I'm only one of many students who appreciates her enthusiasm and encouragement. And Kathleen Currie-Hall has been a model of patient professionalism. Their contributions greatly improved the quality of this work.

First, Eric Bateson showed me what science should look like. Then, he showed me that it can be *both* interesting *and* fun. No small achievement, on either count. If the number of people worth thanking in an acknowledgement is uncountable, then the amount that I've learned from him in the past two years is doubly so. I'm still not exactly sure just what it is he sees in me, but I'm certainly looking forward to what the future holds.

To Laura and Liam.

# **Chapter 1: Introduction**

The act of speaking requires the complex orchestration of myriad neuromuscular components whose mutual action unfolds in space, and just as importantly, through time. At the physical level, the spectrum of speech behavior emerges from the coordinated movement of respiratory, laryngeal, and articulatory musculature. Each moment in the speech stream is an assemblage of the various postures of these interrelated subsystems, each of which operates at different characteristic timescales. For instance, the faster-cycling open/close action of the jaw and the slower inspiration/expiration operation of the lungs must be coordinated simply to sustain the projection of sound. In turn, these quasi-oscillatory rhythms must accommodate the more punctate movements of the articulators to create the fluctuating, patterned waveforms we perceive as speech.

In this way, speech provides a perfect example of the necessity of carefully interlocked timing of events across interrelated subsystems in coordinated biological movement (Turvey 2007). With respect to the contributions of the articulatory and laryngeal subsystems, this notion has received a considerable amount of attention, both theoretically (e.g. Saltzman & Munhall 1989; Browman & Goldstein 1992), and empirically, (e.g. Kelso, Tuller, Vatikiotis-Bateson, & Fowler 1984; Gracco & Abbs 1986; Stone & Vatikiotis-Bateson 1995). However, in spite of the fact that the utterance of even a single CV syllable requires the coordinated movement of the musculature of the entire respiratory air-way system (Abbs & Connor 1989), *which includes* the laryngeal and articulatory systems, respiration has received relatively little attention in efforts to understand the motor coordination of speech. One important outcome of this thesis, which investigates several interrelated effects of varying *vocal effort*, is to underscore the necessity of and set the stage for incorporating respiration into the investigation of speech motor control.

#### 1.1 Vocal effort

Vocal effort can roughly be understood as the level of physiological effort used to produce speech, and as such, has an intrinsic relationship to the acoustic loudness of speech production (Ladefoged and McKinney 1963). However, it is important to emphasize that vocal effort is not strictly equivalent to acoustic loudness. Instead, research has shown that the perception of vocal effort depends crucially on other acoustic factors in addition to loudness, all of which scale directly with intended communication distance (Traunmüller and Eriksson 2000). At the physical level, this involves the functional modulation of the behavior of the neuromuscular subsystems involved in speech production for the purpose of communication under varying conditions of distance and environmental noise (e.g., Lombard speech – Lombard 1911, Pike 1967). Physiologically, it is known to induce robust articulatory (Schulman 1989; Traunmüller 1989; Vatikiotis-Bateson & Ostry 1995), laryngeal (Holmberg et al. 1988; Dromey, Stathopoulos, & Sapienza 1992; Sulter & Wit 1996; Sundberg, Fahlstedt, & Morell 2005; Seshadri & Yegnanarayana 2009), and respiratory effects (Sharp, Goldberg, Druz, & Danon 1975; Iwarsson & Sundberg 1999). In other words, changes of vocal effort with communication distance introduce functionally unique sets of task demands on *all* of the subsystems that enter into speech production, with corresponding adaptations in the structure and time-course of their coordinated activity; that is, vocal effort induces task-specific changes in the character of the various functional synergies (Bernstein 1967; Gelfand, Gurfinkel, Tsetlin, & Shik 1971) mobilized in speech production.

Increased vocal effort is known to induce changes in respiratory patterns on par with those of forced expiration. As the loudness of speech is increased, the speed of inspiratory and expiratory processes is increased (Sharp et al. 1975), a higher percentage of lung vital capacity is expired per breath group (Iwarsson & Sundberg, 1999), and the musculature of the rib cage takes on a primary role in driving the cycle of respiration (Sharp et al. 1975). This includes the action of the intercostal muscles, which are implicated both in respiration and in postural control (Rimmer, 1995). Because these functional changes in respiration have physiological effects that are nearly identical to those observed in forced expiration of air from the lungs (Sharp et al. 1975), we will collectively refer to the inducement of these effects by increasing vocal effort as *vocal forcing*. Taken in combination with the notion that spontaneous speech must be organized with respect to the respiratory cycle (Lenneberg 1967; Chapple 1970; Warner 1979), the presence of these biomechanical linkages suggests that the changes in vocal

forcing should induce functional changes in posture in addition to changes in the speech system proper (Jeong 1991; Vatikiotis-Bateson et al. 2009; Lagier et al. 2010). In turn, these changes must be coordinated in some fashion with the action of the conventional subsystems of speech, as was originally observed by Gould (1971).

Consideration of vocal effort and its modulation of speech breathing patterns also suggests a wider view of the physiological subsystems implicated in speech, beyond the classic tripartite division of respiratory, laryngeal, and articulatory. This idea finds support from the fact that rigid body motion of the head is linked to speech production via a functional coupling with the fundamental frequency of the voice (Yehia, Kuratate, & Vatikiotis-Bateson 2002), and similarly, that the head contributes to speech perception (Munhall, Jones, Callan, Kuratate, & Vatikiotis-Bateson 2004). What's more, the head is also involved in postural control (Dault, Yardley, & Frank 2003). Also, manual gesture is known to be highly coordinated with speech (e.g. Cummins & Port 1996; Leonard & Cummins 2010; Rusiewicz 2010). As with the intercostal muscles of the thorax, the presence of these biomechanical and behavioral couplings suggests that modulation of vocal effort should induce functional changes not only in the coordination of the respiratory, laryngeal, and articulatory subsystems of speech, but also in more remote physiological subsystems coupled to speech production. However, even if vocal effort can be predicted to induce functional changes in coordination across these various subsystems, this raises the question: what form would this coordination take?

#### 1.2 Synchronization in biological systems

Coordinated biological movement emerges as the marshaling of multitudes of neuromuscular components into coherent structures defined by relatively few mechanical *degrees of freedom* (Bernstein 1967), or parameters that specify the actions that a system may take. Bernstein's perspective is complicated by the fact that the actions of living organisms are exhibitions of the intertwined operation of complex physiological rhythms (Glass 1999) which variously converge, diverge, and fluctuate (Winfree 1980).

Von Holst (1939/1973) cast the concept of rhythmic convergence, or *entrainment*, in terms of a general, scale-independent principle termed the *magnet effect*, which takes as its starting point the notion that the actions of biological systems operate at preferred, intrinsic frequencies, or their *maintenance tendency*, as Walsh (1972) later determined to be true of human body segments. Building from this idea of intrinsic or natural frequencies, von Holst went on to show that when multiple neurobiological systems interact, there is a tendency for their intrinsic rhythms to attract one another to their respective preferred frequencies of operation based upon their relative strength (i.e. amplitude). Where one oscillator comes to completely adopt the frequency of operation of another with a specific phase offset, *absolute coordination* is said to take place. This has a direct mathematical analogue in the mode-locking (Pikovsky et al. 2003) that occurs between two out of phase oscillating pendulums connected by a beam, which, over time, will completely synchronize with one another, as Christiaan Huygens first demonstrated in 1649.

However, von Holst recognized that coordinated movement in living systems diverges significantly from this idealized mathematical perspective. What results from the tension between the maintenance tendencies of the individual rhythms and the influence of the magnet effect is a process of *relative coordination* characterized by persistent temporal fluctuations of relative phase. In this state, although the population of oscillators tends toward specific phase and frequency relationships, this tendency is by no means singular. Mode-locking still occurs, but such events are temporary, and the system of coupled oscillators instead explores a range of dynamically stable phase and frequency relationships throughout the time course of behavior.

Even in biological systems, experimental results have shown that the probability of mode-locking increases in the presence of a high amplitude driving force (Schmidt, Shaw & Turvey 1993). With respect to speech, increased strength of vocal forcing should result in an increase of spectral power associated with the respiratory rhythm, the lowest-frequency oscillator of the speech system. This low frequency energy would then be transferred throughout the body, driving the movements of the head and torso at a low resonant frequency. Importantly, movements produced at the same frequency as a resonant

driving force require fewer active degrees of freedom, and therefore the motion can be described more simply at lower dimensionality (Goodman, Riley, Mitra, & Turvey 2000). With respect to vocal effort, this leads to the hypothesis that increased strength of vocal forcing mutually entrains multiple subsystems implicated in speech production; inducing a system-wide state of dynamically stable *relative coordination* characterized by simplification of the collective behavior due to forced oscillation. Specifically, this hypothesis predicts that head motion, torso motion, and speech acoustics should exhibit tighter spatio-temporal coordination as vocal effort increases. This thesis presents evidence that supports this view of vocal effort.

#### **1.3 Postural instability**

Prior research has found a link between increased vocal effort and postural instability (Giovanni, Akl, & Ouaknine 2008), or increases in measures of postural sway. While this effect was attributed to postural stiffening associated with vocal effort (Gould 1971), these researchers did not explicitly address the issue of the conflicting task demands imposed by the needs of postural control and vocalization upon the muscles of the rib cage and the head, which are each implicated in both task domains. Furthermore, with the exception of Vatikiotis-Bateson et al. (2009), no studies have couched the discussion of cross-system correspondence in vocal effort in terms of time-varying coordination, being in the best case dependent upon static analyses using globally calculated correlation coefficients (cf. Lagier et al. 2010). Where the task demands in one domain, speech production, increasingly necessitate use of joint resources shared with another domain, postural control, it stands to reason that the latter will be functionally impaired to some degree, and that this impairment may in fact be related to coordinative processes in the former domain. A major aim of this thesis is to test for this corollary effect, where postural instability is hypothesized to be related to high levels of intra-speaker coordination at high levels of vocal effort.

Postural stability can be construed as a direct physical connection a person has to the environment, insofar as the maintenance of balance is a time-varying process whose goal is the maintenance of a dynamic equilibrium between the physical body and an encompassing physical environment. In other words, maintaining balance is a process that strives to sustain a degree of structural continuity between the body and the environment at the physical level. In the language of coupled oscillatory rhythms, this suggests that postural instability induced by a high level of intra-organism synchronization comes at the cost of this direct environmental connection. Entrainment to a specific rhythm limits the intrinsic structural and rhythmic variability (or flexibility) an organism has at its disposal to adapt to, coordinate with, and navigate its environment.

#### 1.4 Non-stationary time series and fluctuations in the analysis of coordination in biological systems

A major challenge in understanding the temporally conditioned behavior of biological systems is the necessity of analyzing non-stationary time-series (Eke et al. 2000). Living biological systems of all kinds undergo continual change, which implies that their behavior is fundamentally at odds with the assumptions of stationarity that inform many time series analysis methods. Consider the principal requirements of stationarity (Kantz & Schreiber 2004). First, all parameters relevant to a system's dynamics must be held fixed and constant during the measurement period. This applies both to aspects of the experimental setup and the *environment* in which the experiment occurs. More formally, the joint probabilities of finding the system in a given state at one time and another state at another time must be independent of one another during the measurement period. The probability distribution of the measurement cannot change as a function of time. And in a less formally stringent sense, the mean and variance of the time series must be constant throughout the measurement period.

Clearly, these assumptions are fundamentally at odds with the action of anything that could be described as being 'alive'. The condition of constant flux in biological systems is true of both the systeminternal processes that sustain life and, just as importantly, the environmental conditions that a living organism must adapt to and engage with in order to behave. Effectively, then, any and all ecologically valid behaviors, such as spontaneous speech, will be characterized by non-stationary measurements that by definition change over time. The study of biological coordination must take this into account, and make use of methods capable of addressing these issues. A related, yet distinct issue to the appearance of non-stationarity in biological time series is the presence and analysis of *fluctuations*. First, consider a further indicator of non-stationarity discussed by Kantz & Schreiber (2004). If the power spectrum of a signal contains too much power at low frequencies, the time-series must be considered non-stationary, with a particularly interesting case being where power across frequency bands is distributed according to  $f^{-\alpha}$ , that is, according to a power-law. In the case where alpha = 1, the distribution is reflective of 1/f *pink noise* (Bak, Tang, & Wiesenfeld 1987), which appears widely in nature and constitutes an important property of the dynamics of many systems (e.g. Musha & Yamamoto 1997; Ward & Greenwood 2007; Sagues, Sancho, & Garcia-Ojalvo 2007). For example, in the human heart (Peng, Havlin, Stanley, & Goldberger 1995) and respiratory system (Suki 2002; Suki et al. 2003), 1/f fluctuations are an important source of variability associated with healthy function. This is all to say that 1/f fluctuations are not simply random, erratic variation. In time series, this and other types of fluctuation display important structural characteristics in terms of their correlation properties.

Living systems exhibit fluctuations in all manner of their states and behaviors. As such, assessment of coordination within the context of a living system crucially relies on the notion of fluctuation. Quantitatively, fluctuations are related to the variance of a given observable (Bialas & Koch 1999). Similarly, correlations are determined according to the co-variance between multiple quantities, and provide information about the relative co-dependence or independence of the measurements under consideration; in other words, information about the degrees of freedom that characterize the system (Koch 2006). The central problem underlying coordination is the degrees of freedom problem, or how innumerable lower-level, often neuromuscular, components are functionally assembled into a coherent lower-dimensional system (Bernstein 1967). Therefore, in order to understand the underlying dynamics of coordination, it is necessary to characterize both the fluctuations of individual components of a system and, more importantly, the mutual patterning of these fluctuations with respect to one another, all while accounting for the presence of non-stationarity.

In order to address these concerns, the work presented in this thesis makes use of a unique combination of time-series analysis methods in the analysis of the data, including *detrended fluctuation* 

*analysis* (Peng et al. 1995), *correlation map analysis* (Barbosa, Déchaine, Vatikiotis-Bateson, & Yehia 2012) *singular value decomposition analysis*, and *principal component analysis* (Jolliffe 1986; Yehia et al. 2002). This collection of methods is uniquely suited to the task of analyzing coordination in that, in combination, they provide a dynamic view of the interrelationship between subsystem fluctuations, the integration of fluctuations across diverse subsystems in coordination, and the way this mutual patterning is ultimately realized in terms of system dimensionality, which lies at the heart of Bernstein's original (1967) formulation of coordination as the problem of regulating the system's internal degrees of freedom.

#### 1.5 Summary

This thesis reports the results of an empirical study that tests the dual hypothesis that speaking at high levels of vocal effort results in the entrainment of bodily motions associated with speech, and that this entrainment likewise results in postural instability. Data collected from multiple kinematic and acoustic measurement domains were analyzed with a unique combination of methods for the analysis of non-stationary time series. The converging results of these analyses found evidence that increasing vocal effort results in stronger, simplified coordination across a range of subsystems implicated in speech production. Furthermore, evidence of postural instability was found at the highest level of elicited vocal effort. Even though direct measures of respiration (e.g., via RESPITRACE) were unavailable in the present study, this research offers empirical support and a potentially useful methodology for expanding our understanding of the role of respiration in speech production. In addition, studying vocal effort's role in modulating the form of coordination across interlinked bodily subsystems provides a means for examining the complex interrelations between bodily movement and speech acoustics in functional contexts common to everyday use, thus bringing us a little closer to an understanding of speech 'in the laboratory'.

# **Chapter 2: Methods**

In order to evaluate the hypotheses outlined in the previous section, data from an empirical speech production study were collected and analyzed. Spontaneous speech at multiple levels of vocal effort was elicited in the context of a conversational interaction. Measurements of the motion of the head, torso, and the center of pressure (COP) of the net ground reaction forces due to the position of the body's center of mass were collected in addition to speech acoustics. These time-series data were then analyzed using multiple techniques, including *correlation map analysis* (CMA), *detrended fluctuation analysis* (DFA), and *principal component analysis* (PCA). In the remainder of this chapter, details of the experimental methodology, including procedures for data collection and processing, are presented. An in-depth discussion of the analysis techniques appears in Chapter 3.

#### **2.1 Participants**

Seven native English speakers, three males and four females, took part in the study. All were students at the University of British Columbia, aged 20-35. None of the participants were formally trained as actors. All participants were informed of the operation of two of the measurement devices, namely, the OPTOTRAK motion capture system that was used to measure head motion, and the force plates, which were used to measure postural forces and torques at the feet. But despite being informed as to *what* these two devices measure, participants were otherwise naïve to the purpose of the experiment. Data from participants #1-6 were analyzed, with exception of the 'yell' trial for participant #6, who was unable to complete the task. Data from participant #7 were excluded from analysis, due to a persistent error in force plate data collection during the recording session.

#### 2.2 Speaking task and procedures

All participants produced spontaneous speech in the context of an interactive conversation with an investigator. Communicative distance was varied in order to naturally elicit speech at different levels of

vocal effort: a 'normal', comfortable level of speaking at a distance of 1.5m, a 'loud' level at 4m, a 'louder' level at 7.5m, and both the 'shout' and 'yell' levels at 30m. Vocal effort varies along a continuum corresponding to communication distance (Traunmüller and Eriksson 2000). In this way, varying communication distance allows us to perturb the speech system and the body by inducing changes in vocal effort. Coordination is then evaluated by observing and measuring the response to these perturbations.

Data were collected in continuous blocks based on effort condition, with breaks in between changes in vocal effort level. At the beginning of each block, the interlocutor prompted the participant to begin speaking, typically with a question. No instruction as to a specific level of vocal effort was given at the initiation of a block, as the distance between the talker and the interlocutor naturally induced speech at a given effort level. The talkers were simply instructed to speak loud enough so that they could be heard by the interlocutor. An unstructured, spontaneous conversation then ensued between the talker and the interlocutor. Talkers were instructed to drive the interaction, with the investigator prompting with brief questions or interjections as necessary to keep the interaction going. It should also be noted that if the interlocutor had trouble hearing the talker during a trial, they prompted them to 'speak up', as would naturally take place in dyadic speech at a distance. For each of the three lower effort levels, the interaction was first initiated, and four two minute trials were then collected, followed by a break before beginning the next block at a higher level of vocal effort. For the 'shout' level, two two-minute trials were collected, due to the limitations posed by speaker fatigue at this higher level of vocal effort.

In addition to varying communication distance, a 'yell' condition was included, in which participants were instructed to yell angrily at the interlocutor in a fictional scenario, at the same distance as for the 'shout' condition. This task was included in order to examine the effects of a change in discourse style, where limited dyadic response was provided on the part of the interlocutor, and to evaluate whether this change in interaction style would have demonstrable effects on coordination and vocal effort. This task was included in order to corroborate the results of the original pilot analysis, which suggested that postural instability and rigid within speaker coordination occurred in this case. Prior to the beginning of the session, each participant was fully informed that this type of task would be a component of the experiment. Then, each participant was consulted as to a fictional scenario that they would both be comfortable yelling 'in', as well as comfortable performing in the context of a video-recorded experiment, for example, yelling at a referee. While not necessarily ideal from the standpoint of communicative realism, a procedure of this sort was necessary in order to experimentally investigate this type of spontaneous-speech behavior, due to inherent constraints imposed by the laboratory setting. Similar to the 'shout' level, data collection in this condition was limited by the constraint of speaker fatigue. In this case, a single two minute trial was collected for each subject.

In sum, this resulted in a total of 15 trials per subject (4x 'normal' + 4x 'loud' + 4x 'louder' + 2x 'shout' + 1x 'yell').

#### 2.3 Data collection, processing, and reduction

#### 2.3.1 Experimental setup

Each participant completed the speaking task in all effort conditions while standing on a pair of Bertec force plates, one for each foot. A video camera and OPTOTRAK motion capture system were located 2m in front of the participant. Details regarding the collection, processing, and reduction of all collected signals will be discussed subsequently.

#### 2.3.2 Data collection

For all trials, multiple kinematic measurements were collected in addition to speech. This consisted of rigid-body motion of the head, force plate measurements, and 2d motion of the torso extracted from video. Details for each measurement type will be discussed independently.

#### 2.3.2.1 Speech

Speech was recorded using a head mounted Tram-50 interview microphone, located approximately 7.5 inches from the participant's mouth. The microphone was attached to a protrusion of the head-mounted

crown used for collection of rigid body head motion. For each level of vocal effort, the microphone gain was lowered in order to accommodate the louder speaking levels. These signals were later rescaled to a common factor that was determined using a pure sine tone generated at 1 kHz, .23V peak to peak amplitude.

During collection, the speech signal was first fed into a Mackie mixer, split, then taken both directly to video at 48 kHz and simultaneously to a hardware filter, where the speech signal was band-pass filtered at 60-7000 Hz. After A-D conversion, the band-pass filtered speech was recorded to disk using the OPTOTRAK Data Acquisition Unit (ODAU) and NDI first principles software, which facilitated synchronized collection of the filtered speech, head motion, and force plate data.

#### 2.3.2.2 Head motion

Rigid body motion of the head was tracked using the marker-based OPTOTRAK system. Six positional markers were attached to a head-mounted crown worn by the participant. NDI software was used to derive a rigid body representation of the head from the positions of these six individual markers. 6d motion of this rigid body (x, y, z position coordinates in Euclidean space, plus rotations about each coordinate axis) was then recorded to disk at a sampling rate of 60 Hz using NDI first principles software.



Figure 2.1. OPTOTRAK headgear.

#### 2.3.2.3 Torso Motion

Motion of the upper body (Torso) was measured using *optical flow analysis* (OFA – Horn & Schunk 1981, Barbosa et al. 2008) a noninvasive technique for extracting 2d motion from video. First, HD video was recorded at 59.94 frames/sec using a commercial grade camera with intra-frame compression, and recorded to disk with one continuous take for each level of vocal effort. OFA was then applied to a defined region of interest (ROI) within the video frame, for each video file. For torso motion, this region was defined as a rectangle with the top edge at shoulder height, the bottom edge just below the participant's waist, and the outer edges at the most extreme extent of each arm. An example ROI is provided in Figure 2.2 Details regarding subsequent processing are discussed in the next section.





#### 2.3.2.4 Force plates

Two Bertec force plates, one for each foot, were used to measure ground reaction forces and their associated torques. Each force plate provides a six component signal, consisting of voltage measurements indexing forces in three translation dimensions (x, y, and z in 3d Euclidean space), and the associated moment (torque) about each of these three axes. These force and torque components will subsequently be referred to as Fx, Fy, Fz, (forces in each direction) and Mx, My, Mz (moments about each axis). The force plate data were also collected using the ODAU at a sampling rate of 7000 Hz. Refer to Figure 2.3 for the orientation of the coordinate system relative to the experimental setup. The force plates were

zeroed with the participant already standing on them, necessitating the addition of a correction term during data processing in order to yield the true values for the force measured in the vertical (Fz) direction.



**Figure 2.3 Bertec Force plates with orientation axes.** Force measured along the z axis is positive in the downward direction. Force measured in the y direction corresponds to the anteroposterior axis, force in the x direction corresponds to the mediolateral axis. Only the moment about the z-axis is pictured.

#### 2.3.3 Data processing

All data processing, reduction, and analysis was performed with MATLAB. All signals were resampled to 60 Hz prior to analysis. As mentioned previously, the speech, head motion, and force plate data were collected simultaneously using NDI first principles software and the ODAU, so no further steps were required to align these signals.

In the case of the video based data, an additional processing step was required in order to ensure proper alignment with the ODAU data. This was achieved by aligning the audio that was recorded directly to video with the hardware filtered audio collected by the ODAU using a correlation function. This provided the video frame numbers associated with the stop and start of each trial. These frame numbers were then used to extract the portion of the optical flow signal corresponding to each trial from the entire optical flow signal (see section 2.3.2.3) computed for each video. This 59.94 Hz signal was then converted to 60 Hz to complete the alignment process.

The raw voltage measurements of the force plates were first down-sampled to 60 Hz, and then converted to force and torque values through multiplication by appropriate scale factors (Bertec Force plate Manual, 2009). A correction term was then added to the measured values for the force in the z direction. This correction term was computed according to Equation 2.1,

$$F_{zc} = ma, \tag{2.1}$$

where  $F_{zc}$  is the correction term, *m* the estimated mass of the participant in kg, and *a* the acceleration due to gravity, 9.81 m/s<sup>2</sup>.

#### 2.3.4 Data Reduction

#### 2.3.4.1 Speech—F0, acoustic RMS

F0 contours were extracted from the 7 kHz hardware filtered audio using YAAPT (Zoharian & Hu 2008). YAAPT's parameters were set with an F0 search range of 60-550 Hz in order to capture the extended range of F0 values known to accompany increases in vocal effort (Traunmüller and Eriksson 2000, Lienard and di Benedetto 1999). Following extraction, these pitch contours were then downsampled to 60 Hz for subsequent analysis.

Acoustic root-mean-square (aRMS) was also computed directly from the hardware filtered audio in order to provide a time-varying measure of acoustic amplitude. This was done using a frame-based technique (Barbosa, Yehia, & Vatikiotis-Bateson 2007) which directly aligns the speech audio to the target sampling rate (one audio frame per target motion rate sample), in this case, 60 Hz.

#### 2.3.4.2 Rigid body head motion

The 6d rigid body motion of the head was reduced to a scalar time-series using a Euclidean distance metric on the x, y, and z coordinates for each time sample, providing a uni-dimensional representation of the displacement of the head (in mm) through time.

#### 2.3.4.3 Torso motion

The optical flow algorithm provides vector values and magnitudes in both the x and y directions for each video frame. In addition, a combined sum in both directions is computed, which provides a measure of the total magnitude of motion within the region of interest for a given frame of video. In the analysis, the combined sum, and therefore a scalar measure of the total amount of motion in the region of interest for each video frame, was used.

#### 2.3.4.4 Force plates —center of pressure (COP)

The center of pressure is a measure of the centroid position of the ground reaction force due to all downward forces acting on the force plate. These forces are due to the various neuromuscular systems responsible for postural control and balance (Winter 1995).

The center of pressure for each individual plate was first calculated using the respective force and torque measurements (Bertec Force plate Manual 2009). A single, global center of pressure, representing a combination of the individual centers of pressure for plates one and two, was then calculated using a change of coordinate system. For analysis, these 2d time-series were then converted to a uni-dimensional scalar representation using a Euclidean distance metric, analogous to the method used for the reduction of the head motion data. This yielded displacement time-series for the individual plate COPs, and the global COP (in m).

F0
aRMS
Optical Flow: total magnitude of motion
Head: displacement
COP - global: displacement
COP - plate 1: displacement
COP - plate 2: displacement

Table 2.1. List of 60 Hz time series obtained after post processing and data reduction

## **Chapter 3: Analysis & Results**

Several interrelated techniques for the analysis of non-stationary time series were employed in the analysis of the data. Correlation map analysis (CMA) was used to quantify the dynamic structure of correlations between the time series listed in Table 2.1, that is, to investigate their time-varying correspondence. The relative complexity of the temporal structure of correlations was then evaluated with a widely used dimensionality reduction technique, singular value decomposition (SVD). These methods were combined with detrended fluctuation analysis (DFA) of the individual signals, which quantifies the nature of signal fluctuations in terms of their degree of statistical self-affinity and spectral characteristics. Finally, principal component analysis (PCA) was used to evaluate the degree of redundancy present in each of the original 6d force plate signals, providing a measure of the co-dependence between their component forces and torques across different levels of vocal effort. A description of each of these techniques immediately follows. These descriptions are then followed by the results of the analysis.

#### **3.1Analysis Techniques**

#### 3.1.1 Correlation Map Analysis (CMA)

Biological behaviors are the sum-total output of the mutual interaction of myriad biological rhythms operating across multiple spatial and temporal scales, and speech is no exception. CMA (Barbosa et al. 2012) provides a means of quantifying these patterns of time-varying coordination between pairs of signals by calculating the instantaneous cross-correlation between them as a function of both time and temporal offset, according to Equation 3.1,

$$\rho(k) = \frac{S_{xy}(k)}{\sqrt{S_{xx}(k)S_{yy}(k)}},$$
(3.1)

where  $\rho(k)$  is the instantaneous correlation coefficient,  $S_{xy}$  is the instantaneous covariance between signals x(k) and y(k) centered at any point  $k = k_0$ , and  $S_{xx}$  and  $S_{yy}$  are the instantaneous autocovariance of signals x(k) and y(k), respectively. This process yields a two-dimensional correlation map which provides a

quantitative representation and visualization of the instantaneous correlation between the two input signals across time and a desired range of temporal offsets. See Figure 3.1 for an example. The 2D method enables the quantification of fluctuating changes in the phase relationships of correlations with the passage of time, as often occur in speech (Gracco & Abbs 1986) and other biological behaviors where strictly time-locked synchronization rarely takes place (see von Holst 1939/1973).



**Figure 3.1 Example correlation map.** Head motion and acoustic RMS. Instantaneous cross-correlation between the two signals is calculated as a function of time (ordinate) and temporal offset (abscissa). The color scale on the right indexes the instantaneous correlation values  $\rho(k)$ . Perfect positive correlations  $\rho = 1$  (in-phase) are red, while perfect negative correlations  $\rho = -1$  (anti-phase) are blue. The cursor at approximately 82 seconds illustrates an offset relation between the two signals. The positive correlation at this 1(s) offset position is the result of the computation of the instantaneous cross-correlation centered at the positions indicated by the vertical bars in the signal panes, where the head motion signal is leading aRMS by one second, and both signals are clearly trending upward. Note that this strong positive correlation is in addition to the one clearly visible at approximately 0(s) offset.

In the present experiment, CMA was used to quantify the time-varying patterns of correlation among the kinematic and acoustic measures listed in Table 2.1. In total, eleven pairs of signals were analyzed for each participant. This consisted of an exhaustive pairwise matching of five signals: F0, aRMS, head motion, torso motion, and global COP displacement, yielding ten pairs, plus an additional pairing of the COP time-series calculated for the individual force plates. For analysis, separate cross-correlograms for positive and negative correlations were constructed using the following procedures. First, for each signal pair in a given trial, the number of significant instances of correlation (either rho > .5 or rho < .5) are totaled individually for each temporal offset. This produces a histogram consisting of the number of instances of significant correlation, or 'hits' at each offset. These histograms are then normalized to the length of that trial. This normalization effectively converts the total hits measure for a given offset into the percentage-total instances of significant correlation for each offset (%COR). In practice, the normalized cross-correlograms enable the determination of the predominant in-phase and anti-phase offsets between the signals of interest. Also, variability in the range of offsets and the strength of correlation for the purposes of comparison across effort conditions, the mean of the percent-total correlation (m%COR) across all offsets for each of the positive (m%COR-pos) and negative (m%COR-neg) cross-correlograms was selected as the principal measure of interest for subsequent analysis.



**Figure 3.2 Correlation map and cross-correlograms**. F0 and acoustic RMS. The positive cross-correlogram (upper right) shows a strong peak at approximately zero offset, corresponding to the band of positive correlation along zero-offset in the correlation map.

#### 3.1.2 Singular value decomposition (SVD)

If increased vocal effort results in postural entrainment, a systematic simplification of the dynamics of coordination between the subsystems under investigation will take place. In particular, the intermittent rhythms of the head and torso will become increasingly coactive with the continuous rhythms of speech at higher levels of vocal effort, resulting in reduced variation of the phase relations between these quasi-periodic oscillatory systems. In the correlation map, such a reduction in time-varying complexity would be evidenced by greater saturation of correlations at particular offsets, as defined by the periodicities associated with the combined dynamics of the involved subsystems, and driven by the periodic forcing of the respiratory cycle. Re-cast in the terminology of the degrees of freedom problem presented by Bernstein (1967), this can be construed as dimensionality reduction in the temporal domain. Movements at resonance with a coupled periodic forcing function are known to exhibit dimensionality reduction as compared to those at frequencies other than resonance (Goodman et al. 2000). If the forcing of the respiratory cycle at high levels of vocal effort is sufficient to drive the system into resonance, dimensionality reduction in correlation structure should take place across multiple measurement domains, including kinematic-kinematic and acoustic-kinematic pairings, in addition to pairing of the kinematics and acoustics with the postural dynamics measured by the force plates.

In order to assess whether or not this effect occurs, singular value decomposition was performed on the correlation maps. SVD is a matrix factorization technique that can be used to identify lowerdimensional underlying structure in high dimensional data. The factorization leads to a representation of the original dataset X in terms of two orthogonal basis matrices U and V, and a diagonal matrix  $\Sigma$ consisting of the singular values, according to Equation 3.2,

$$X = U\Sigma V^T, (3.2)$$

where the superscript <sup>T</sup> denotes the matrix transpose. From this representation, the original data can be approximated in terms of the components of the basis matrices U and V that contribute most to the underlying structure of the data, as indexed by their associated singular values.

In an intuitive sense, the singular values correspond to the relative 'strength' of each of these components in representing the variance of the original data. Note that this is similar to the ranking of component contribution to variance provided by PCA discussed in section 3.1.4 below (see Shlens 2014 for further discussion of the relationship between PCA and SVD). Highly redundant data, in which a strong underlying pattern exists, will require fewer components to approximate the original than unstructured data, which will require many more components for its approximation. For a general discussion of SVD's application to dimensionality reduction, see Rajaraman & Ullman (2012).

In the correlation maps, this notion of underlying structure is reflected in the correspondence of correlations across offsets and through time. Maps whose structure is due to a simple underlying pattern will display greater consistency of correlation structure across offsets, and thus, reduced dimensionality as compared to those whose structure is more complex, or even random. This underlying pattern is determined by at least two factors: (1) temporal patterning of the signals with respect to one another, and (2) the simplicity, or *smoothness*, of the signals themselves. Figure 3.3 provides several examples using idealized signals, including sine waves (3a), (3b), (3c), brown noise (3d), pink noise (3e), and white noise (3f).

In Figure 3.3, the maps generated by signals with simpler underlying dynamics (e.g. sine waves) exhibit highly regular correlation patterns. Saturated bands of high correlation, both positive and negative, persist across all lags and through time. In the random walks, these bands are intermittent and display less structural consistency. In the noise map, they are nonexistent. This qualitative difference between signal types can be quantified by the SVD analysis in terms of number of singular values necessary to represent the variance of the original data at a desired level of accuracy. Figure 3.4 provides an example. The cumulative variance explained is plotted as a function of the number of singular values for each example map. Maps with coherent structure require fewer singular values to reconstruct the original data with 99.8% of the variance explained. The maximum number of singular values possible is equivalent to the number of offsets in the correlation map, in this case, 301.



**Figure 3.3 Correlation maps for idealized signals.** Sample correlation maps for (a) sinusoids with identical frequencies (2 Hz), (b) sinusoids with different frequencies (1 Hz, 2 Hz), (c) equal frequency sine waves, one with added noise, (d) brown noise (i.e. two random walks), (e) pink noise, and (f) white noise. All signals sampled at 60 Hz for n=1000 samples  $\approx$  16.67 seconds. Visible window is the range from 2 seconds to 14 seconds. Offset range is +/- 2.5 seconds, which equates to 301 possible offsets.


**Figure 3.4. SVD of idealized correlation maps**. Cumulative variance plot for SVD of the correlation maps in Figure 3.3. Compared to the number of components (singular values) necessary to represent the Gaussian white noise map with 99.8% accuracy (144 SVs), fewer components are needed for the equal frequency sine wave map (7), equal sine waves, one with added white noise (19), mismatched frequency sine waves (22), brown noise (41), and pink noise (103). Note the influence of the spectral properties of the signals on the structure of the correlation map. Maps generated with signals whose spectra are characterized by increased relative concentration of power in fewer frequency bands require fewer singular values for their reconstruction.

## 3.1.3 Detrended Fluctuation Analysis (DFA)

An issue related to dimensionality reduction in the context of movements at resonance is the change in the distribution of spectral power due to the influence of the driving force. In this case, it is hypothesized that the increased strength of the low-frequency driving force of the respiratory cycle will result in a concomitant increase in low-frequency spectral energy, and increased 'smoothness'/simplicity, in the signals of interest.

The smoothness of a signal is related to the type and amount of variation, or *fluctuation* in its spatial and temporal characteristics. For example, a sine wave follows a strictly deterministic path, and therefore does not exhibit fluctuations. White noise, on the other hand, consists entirely of fluctuations

about some mean value, while pink noise exhibits a structured subset of the fluctuations seen in white noise. In terms of spectral characteristics, the spectrum of a simple sine wave will have all power concentrated at a single frequency, power will be distributed equally across all frequencies in white noise, and in pink noise power across frequency bands is distributed according to 1/f, that is, with higher relative concentration of spectral power in lower frequency bands.

DFA provides a means of quantifying this idea in terms of a single quantity, the scaling exponent  $\alpha$ , which is an adaptation of the generalized Hurst exponent for non-stationary signals (Peng et al. 1995). For a given time-series,  $\alpha$  indexes the presence or absence of long-timescale temporal correlations. Correlation timescales relate directly to the structure of the power spectrum. In a rough sense, long-timescale correlations are due to low frequency spectral components present in the signal, and shorter timescale correlations relate to higher frequency spectral components present in the signal. Furthermore, information regarding the relative distribution of correlation timescales evident in a signal's fluctuations is contained in the value of  $\alpha$ , providing a means of assessing the presence of statistical self-similarity in the structure of the signal's fluctuations. As such, the scaling exponent has a straightforward interpretation in terms of its relationship to the structure of the power spectrum associated with the color of noise corresponding to its value (Buldyrev et al. 1995).

Calculation of alpha is as follows. First, the time-series x(t) is integrated according to Equation 3.3,

$$y(n) = \sum_{t=1}^{n} x(t).$$
 (3.3)

Then, the integrated time-series y(n) is subdivided into windows of length L. For each window, a local least-squares trend is fit to the data, according to Equation 3.4,

$$argmin_{a,b} \left\{ E^2 = \sum_{n=1}^{L} (y(n) - an - b)^2 \right\}.$$
(3.4)

The slope and intercept parameters of the linear fit, *a* and *b*, are minimized with respect to the squared error,  $E^2$ . Then, the RMS deviation from these fitted trends, or fluctuation *F*(*L*), is calculated for all window lengths *L*, as in Equation 3.5,

$$F(L) = \left[\frac{1}{L}\sum_{n=1}^{L}(y(n) - an - b)^2\right]^{1/2}.$$
(3.5)

This process is repeated for multiple window lengths (timescales). In the implementation by Little, McSharry, Moroz, & Roberts (2006), the number of timescales is determined according to  $n = log_2(M)$ , where M is the length of the time-series. Each timescale is then determined as the number of subdivisions according to powers of two (eg.  $2^0 = 1$  subdivision,  $2^1 = 2$  subdivisions ... $2^{n-1} = n$  subdivisions), up to the total number of timescales. The scaling exponent  $\alpha$  is then found as the slope of a linear fit on the graph of *F*(*L*) against *L* on a log-log scale.

Interpretation of the scaling exponent is as follows. Values of  $\alpha \approx 0.5$  correspond to white noise. In white noise, fluctuations from one moment to the next are completely independent of one another, or in other words, uncorrelated. For values  $0.5 < \alpha \le 1$ , long term correlations are present in the fluctuations. In this range, the distribution of correlation time-scales follows a power-law form, with  $\alpha \approx 1$  being the special case of 1/f pink noise. In this case, correlations exist across all timescales, with the relative frequency of correlations decreasing with increasing timescale length. Importantly, a signal whose fluctuations exhibit a power law distribution of correlation timescales is said to have no single characteristic timescale that dominates the dynamics of the process. This contrasts with the behavior of fluctuations whose scaling exponent lies in the range  $1 < \alpha < 1.5$ , with  $\alpha \approx 1.5$  indicating brown noise. In this range, the fluctuations exhibit long-term correlations, but the distribution of time-scales does not follow a power law. Instead, the signal has a characteristic long-period timescale, with the relative dominance of low-frequency fluctuations resulting in a smoother signal profile.

This gradation from  $\alpha = 0.5$ -1.5 can also be interpreted as a cline that represents the smoothness of the signal, as shown in Figure 3.5. Larger values of the scaling parameter indicate a smoother signal whose spectrum contains increased concentration of spectral power at lower frequencies. On the other hand, values closer to 0.5 correspond to the 'roughness' of white noise (Peng et al. 1995), with power distributed equally across all frequency bands.



**Figure 3.5 DFA of example noise signals**. The white, pink, and brown noise signals used to generate the correlation maps in Figure 3.3 are displayed with the value of the scaling exponent computed by DFA. The brown noise signals are smoother than the pink and white noise signals, owing to the increased magnitude of long time-scale fluctuations relative to short time-scale (i.e. low vs. high frequency) fluctuations.

The power-law range of  $.5 < \alpha < 1$  is a special case that merits further explanation. Power law distributions occur in a wide variety of biological systems, including the structure of airways in the human lung (Suki et al. 2003), and in the distribution of fluctuations in inter-breath interval lengths in respiration (Frey, Silverman, Barabasi, & Suki 1998). These distributions have several important properties, but the crucial distinction between them and other probability distributions occurs with respect to the concept of *scale*. For many distributions, the mean value of a measurement reflects the 'typical' size or scale of the quantity of interest, as for example, in the distribution of human heights in a large population (Newman 2006). But power law distributions tend to vary across an extremely wide range of values, exhibiting the so-called "long tail". They lack a typical value, and instead, the values of the

distribution for any given value of the measured variable are proportional to one another. In the simplest case, this is expressed in general form as a scaling relation, given by Equation 3.6,

$$D(x) = Cx^{-\beta}, \tag{3.6}$$

where D(x) is the value of the distribution function (e.g. the probability mass function) at some value *x*, *C* is a constant, and  $\beta$  is the scaling parameter. For instance, in the case where C = 1 and  $\beta$  = 1 (pink noise), values of x = 1 would occur twice as often as values of x = 2, and three times as often as values of x = 3 (i.e. D(1) = 1 vs. D(2) = 1/2, vs. D(3) = 1/3). Taking the logarithm of both sides of Equation 3.7 yields

$$\log D(x) = -\beta \log x + \log C, \tag{3.7}$$

which shows that values of the distribution function D(x) are linearly related to values of the independent variable *x* on a log-log scale, with the slope of the linear relation given by the scaling parameter  $\beta$ . The presence of this kind of *power law scaling* within a distribution implies that there is no typical value that characterizes the distribution. Instead, the value of D(x) shows a strong dependency on the value of x, and importantly, these measures are directly proportional to one another across a range of scales.



**Figure 3.6. DFA of signal with power law fluctuations.** The 60 Hz signal (top) exhibits power-law fluctuations indicated by the scaling exponent  $\alpha = .939$ , as determined by the slope of the linear fit of F(L) ~ L on a log-log scale. Fluctuations calculated for short temporal intervals are proportional to fluctuations calculated over long temporal intervals.

This notion of proportionality across scales extends directly to the idea of mathematical *self-similarity*. This idea is most commonly expressed in geometric terms. Geometrically self-similar objects are known as *fractals* (Mandlebrot 1977). The pieces of a self-similar fractal object are constructed of smaller copies of the whole object. This part-whole similarity is revealed by viewing the pieces of the object under successive levels of magnification, and is defined in terms of proportional scaling. In fact, power-law scaling is a special case of mathematical self-similarity, and 1/f power spectra are characteristic of self-similarity (Bassingthwaighte, Liebovitch, & West 1994). Mathematical self-similarity can also apply to objects or processes whose pieces resemble the whole in terms of their

statistical properties. Unsurprisingly, this is referred to as *statistical self-similarity*. In statistically selfsimilar objects or processes, the scaling relationship between pieces of different size is construed in terms of the relationship between a measured property and the scale, or resolution, at which the property is measured, according to Equation 3.8,

$$M(as) = kM(s), \tag{3.8}$$

where M(s) is the property measured at a scale *s*, k is a constant of proportionality, and *a* < 1 is a parameter that reduces the measured scale. For a power law, this equation can be rewritten as in Equation 3.9,

$$M(as) = M(s)a^{\alpha}, \tag{3.9}$$

where alpha is the power law scaling exponent. Equation 3.9 states that for some measure M(s) at a scale s, a measure of the same object at a smaller scale M(as) is related to the measure of M(s) by the decrease of measurement scale a raised to a scaling exponent  $\alpha$ . The measured property of smaller pieces of the whole object are similar to the measured property calculated for the whole object. This sense of measure is generic, and can apply to a wide variety of properties. For instance, the magnitude of fluctuation F(L), measured by DFA, as shown in the example provided by Figure 3.6.

With respect to time series, statistical self-similarity refers to similarity in the statistical properties of shorter intervals of the series with respect to larger intervals, or even the entire time series. The relation between properties calculated at different temporal scales can be expressed as a functional scaling law similar to Equation 3.9. If we let  $a = 1/\lambda$ , where  $\lambda$  defines the number of subintervals of the time series, then Equation 3.9 becomes,

$$M\left(\frac{s}{\lambda}\right) = M(s)\lambda^{-\alpha}.$$
(3.10)

Defining M(s) as the probability mass function of the time series x(t) defined over the interval t = 1, ..., N, we obtain,

$$D(x(t)_1, \dots, x(t)_{\frac{N}{\lambda}}) = \lambda^{-\alpha} D(x(t)_1, \dots, x(t)_N), \qquad (3.11)$$

where  $D(x(t)_1, ..., x(t)_N)$  is the probability mass function of the time series x(t) measured over the entire interval, and  $D(x(t)_1, ..., x(t)_{N\lambda})$  is the probability mass function over subintervals of the time series of length  $N/\lambda$ . This equation states that the distribution of the time series over subintervals whose length is  $N/\lambda$  is given by the distribution of the entire series multiplied by a factor. In this way, the statistical properties of shorter intervals are proportional to the properties of the entire time series. This effect can be visualized in the self-similarity of histograms calculated over different time intervals, as shown in Figure 3.7.



**Figure 3.7. Statistical self-similarity of time series with power law fluctuations**. The idealized time series from Figure 3.6 with scaling exponent  $\alpha = .939$  viewed at successively higher levels of temporal resolution (left). Note the qualitative similarity of the signal under each level of magnification. The statistical self-similarity of the signal's structure is reflected in the similarity of the histograms of squared values for the signal  $x(t)^2$  calculated for successively smaller temporal intervals (right). Each interval length is equal to N/ $\lambda$ , where N is the length of the signal in samples. The number of hits in the histograms (y-axis) scales as  $\lambda^{-\alpha}$ .

Power-law scaling and self-similarity in the statistical characteristics of a signal has important consequences for the behavior of the system in question. For example, 1/f power-law fluctuations are indicative of healthy variability in physiological systems such as the human heart (Musha & Yamamoto 1997; Peng et al. 1995) and respiratory system (e.g. Suki 2002). In the present study, a particularly important property of self-similar processes is their greater adaptability to internal changes and to changes

in the environment (West 1990). This study investigates coordination among several interlinked heterogeneous subsystems in the context of speech production. The change induced by increasing respiratory forcing with increased vocal effort is thought to have important consequences for this coordination. In this sense, the response characteristics of these heterogeneous systems should exhibit some common changes in their fluctuations due to the respiratory perturbation. However, the fluctuation properties of these individual components will also likely display intrinsic differences; movements of the torso could tend to show power law fluctuations while the movement of the head could instead occupy a smoother range on the gradient indexed by the DFA scaling exponent, i.e.  $1 < \alpha$ . In this case, the differences in fluctuation structure could ultimately influence coordination, as they would index differences in the ability of the individual subsystems to adapt to the respiratory perturbation.

DFA was performed on the signals listed in Table 2.1 and the individual force and torque components (Fx, Fy, Fz, Mx, My, Mz) of both force plates, using the implementation in Little et al. (2006). For each condition, signals of a given type were concatenated across trials and analyzed as one contiguous signal, resulting in a single scaling exponent measurement per signal type for each participant in each condition.

### 3.1.4 Principal Component Analysis (PCA)

It is hypothesized that the stability of correspondence among the component variables of the individual force plate measurements (x, y, z, forces and torques) will diminish at the highest levels of vocal effort. This was assessed using principal component analysis, which in recent years has found increasing application in studies of motor control and coordination (Daffertshofer, Lamoth, Meijer, & Beek 2004). Like SVD, PCA is a technique for finding a low-dimensional representation of high dimensional data in terms of its underlying internal structure. In fact, the two techniques share a close mathematical relationship, in that PCA can be derived directly from the SVD (see, for example, Shlens 2014, Yehia, Rubin & Vatikiotis-Bateson 1998; Yehia, Kuratate, & Vatikiotis-Bateson 1999). Briefly, given a data matrix X, a matrix Y is created according to Equation 3.12,

$$Y = \frac{1}{\sqrt{n}} X^T, \tag{3.12}$$

where n is the number of observations of the variables in X. Then, the SVD of Y yields the two changeof-basis matrices and the diagonal matrix consisting of the singular values. The columns of the changeof-basis matrix V are the principal components of X,

$$Y = U\Sigma V^T. ag{3.13}$$

In an intuitive sense, PCA begins with the assumption that the variances in a multivariate dataset are representative of important patterns in the data. In the case of time-varying data, this corresponds to important system dynamics. Building from this notion, the variables of the original dataset are transformed in such a way that the variance of the original data can be accurately represented by a linear combination of uncorrelated variables, the principal components. This process enables structural redundancies in the original data, which manifest as high co-variances, to be compactly represented by fewer variables in the principal component representation.

The PCA decomposition is achieved by first transforming the data using a change of coordinate system. The new coordinate axes are determined iteratively. The first principal component (i.e. coordinate axis) is selected as the direction maximally associated with the variance, or 'spread', of the data. The data are then projected onto this new coordinate axis, and then subtracted from the original. This 'removes' the variance associated with this principal component. After this subtraction, the next principal component is determined according to the same principles, but with each successive component being subject to the constraint that it be orthogonal to the previous component.

Two examples, the first trivial and the second less so, will be helpful in motivating this description of PCA. First, consider the example in Figure 3.8. The three original variables are identical 2 Hz sinusoids. Both the time-series representation of the individual variables and their plot in three dimensional space are shown. In 3d space, note that the data are spread along a single straight line. This has two important implications: i) the data are perfectly correlated, or in other words, redundant, and ii)

the important dynamics of the system can be described in terms of one dimension (i.e. one variable) rather than three.



Figure 3.8 PCA example one: Three 2 Hz sinusoids. Plot of the original variables in 3d space, plus timeseries plots of both the original variables and the principal components.

The first principal component lies along the direction of the spread of the data. Because all data points lie exactly along this dimension, the process of re-projection and subtraction removes 100% of the variance, and no further principal components are necessary to represent the data. In other words, the first principal component 'explains' 100% of the variance, as indicated in Table 3.1. That the data can be reduced to a single 2 Hz sinusoid is reflected in the principal component time-series. In addition, the contributions of the original variables to the variance accounted for by the first principal component can be determined. In this case, all variables contribute equally to the first principal component.

Table 3.1 PCA example one. Percent variance explained by each principal component and variance contributed
by each variable to the first principal component in the 2 Hz sinusoid example shown in Figure 3.8.

Variable #	PC 1: % Variance contributed		
1	33.3		
2	33.3		
3	33.3		
PC#	% Variance explained		
1	100		
2	0		
3	0		

Next, consider a slightly more complicated example consisting of three 2 Hz sinusoids, one with added white noise and one with added pink noise. The fact that noise obfuscates the common sinusoidal pattern in the dataset is reflected in the 3d scatter plot in Figure 3.9(a). Although there is a dominant spread to the data that reflects the correlation among variables due to the underlying sinusoidal process, the data are also spread along other dimensions due to the added noise.

The length of the arrows reflects the strength of each principal component in explaining the variance of the data. In technical terms, the vector for each component is scaled according to the eigenvalue associated with each principal component, which are computed as eigenvectors of the covariance matrix of the data (Shlens 2014). Analysis of the relative strength of the individual principle components in explaining the variance of the data is alternatively referred to as analysis of the *eigenvalue spectrum*. In the present example, not only does the first principal component lie along the direction of greatest variance, but a large majority of the variance of the data is associated with this principal direction relative to that explained by the other principal components, as Table 3.2 shows. The orthogonality of principal components two and three is evident in the expanded side view in (b).



**Figure 3.9 PCA Example two: Three 2 Hz sinusoids, two with added noise**. Original data plotted in three-dimensional space with principal component axes.

Similar to the previous example, the common underlying sinusoidal pattern is captured by the first principal component, as shown in Figure 3.10. However by looking at the original variables, we see that the noise present in variables two and three can be thought of as constituting a secondary source of redundancy. Hence, the first principal component resembles a 2 Hz sinusoid with an element of added noise. This is evident in the contributions of the individual variables, or their *loadings*, to the first principal component listed in Table 3.2. Both sinusoids with added noise contribute more variance to its construction, although the contribution from the sinusoid with added white noise is considerably stronger. Principal components two and three each explain the remainder of the variance due to the added noise. The strength of each is reflected in the variance of the time-series pictured in Figure 3.10, with principal component two having a higher variance,  $\sigma^2 = 0.218$ , than principal component three,  $\sigma^2 = 0.0714$ .



Figure 3.10 PCA example two: Original variables and principal components.

**Table 3.2 PCA example two.** Percent variance explained by each principal component and percent variance contributed by each original variable to the first principal component for the noise-contaminated sinusoids example.

PC#	% Variance explained		
1	84.5		
2	11.7		
3	3.8		
Variable	PC 1: % variance contributed		
Variable sine 2 Hz	PC 1: % variance contributed 28.5		
Variable sine 2 Hz + w-noise	PC 1: % variance contributed 28.5 40.1		
Variable sine 2 Hz + w-noise + p-noise	PC 1: % variance contributed 28.5 40.1 31.4		

Contrasting these two examples, we see that more redundancy in the data—i.e., better correspondence between the component variables—requires fewer principal components to fully represent the data. In the first example, the perfect correspondence between the original variables results in a reduction to a single principal component that explains 100% of the variance of the data. In the second example, representing 100% of the variance necessitates all three principal components. However, the first two principal components are sufficient to accurately represent the structure of both the

sinusoidal process and the added noise<sup>1</sup>, with over 96% of the variance explained between them. In this way, the important dynamics of the three-dimensional system are said to be reducible to a two-dimensional representation.

With respect to the present experiment, changes of vocal effort are expected to affect the correspondence among the component forces (Fx, Fy, Fz) and torques (Mx, My, Mz) of the 6d force plate signals. High vocal effort is expected to result in decreased correspondence between these component variables, which will necessitate the inclusion of more principal components for the representation of the data at the desired level of accuracy.

#### 3.1.5 Statistical Analysis

Following Levene's test of homogeneity of variance, one-way within subjects ANOVA (SS = Type III) was performed on the results of the CMA, DFA, and PCA. This was done in order to examine the effect of effort condition while controlling for variability across-subjects and to account for the limitations of the unbalanced experimental design, due to different numbers of trials being recorded for different levels of vocal effort. Because a within-subjects analysis was performed, effect size was calculated as partial omega squared,  $\omega_p^2$ , which offers the additional benefit of reducing bias due to small sample size (Lakens 2013). Post-hoc tests were performed with Tukey's HSD, which provides a conservative estimate of between condition differences. In order to simplify the in-text presentation, post-hoc test statistics are contained in the appendices.

## **3.2** Effects of vocal effort on acoustic signal magnitudes

In order to verify that increasing communication distance properly indexes increased vocal effort, analysis of the magnitude of the acoustic RMS (aRMS) and F0 signals was performed. As expected, the average

<sup>&</sup>lt;sup>1</sup> Here it is worth re-emphasizing that oftentimes noise is not simply disordered 'randomness', but instead consists of fluctuations that contribute to the appearance of spatio-temporal order in a wide variety of systems (see Sagues et al. 2007 for a review).

(RMS) amplitude of aRMS was higher in each successive vocal effort condition for all participants, with the exception of participant five, whose peak was in the 'shout' condition, despite having performed the task correctly. Figure 3.11 provides a summary of this information. In the figure, error bars have been omitted in order to aid visual clarity, as the variability across trials within a given condition (i.e. at a given distance) was low for each participant. Note that scales vary across participants due to the use of different recording levels for each session, while the scale within participants was determined with respect to a common scale factor. This scale factor was recovered by playing a pure sine tone at 1 kHz, .23v peak-to-peak amplitude at the beginning of the first trial of each condition block. Amplitudes across effort conditions were then scaled relative to the recovered amplitude of this sine tone in the 'normal' effort condition for each participant.



**Figure 3.11. Average magnitude of acoustic RMS for individual participants**. Calculated as RMS(aRMS). Mean amplitudes by condition. Scale across participants varies due to the use of different recording levels for each participant. Scales within participants are adjusted to a common factor determined by a pure sine tone @ 1 kHz .23V peak to peak amplitude.

Comparing the effects of vocal effort condition on aRMS amplitude across participants, one sees that although the increase of aRMS is systematic across successive levels of vocal effort, the precise degree of this increase varies across participants. Generally, a 50%-100% increase in aRMS takes place in the change from 'normal' to 'loud', 50%-100% again from 'loud' to 'louder', a four-fold increase in the transition from 'louder' to 'shout', and anywhere from 50% to 300% in the final switch to 'yell', excluding participants five and six. By and large, these findings are in agreement with analogous results reported for SPL changes as a function of increased communication distance (Traunmüller & Eriksson 2000; Lienard & di Benedetto 1999).

F0 has also been reported to increase with increased vocal effort (eg. Tranumuller & Eriksson 2000; Lienard & di Benedetto 1999). This effect was also observed in the present study, as summarized in Figure 3.12. Average (RMS) magnitude of F0 was computed for each trial, taking only voiced speech segments into account. Again, the expected systematic tendency of increasing F0 with increased vocal effort is observed for all participants, excluding participant 5 in the 'yell' condition. On the other hand, participant 4's F0 increase in 'yell' was marginal.



Figure 3.12. Average Magnitude of F0 (Hz), individual participants. Effort condition means, voiced speech segments only.

## 3.3 Vocal effort and breath group period length

An estimation of the length of the breath cycle period was determined in order to define a window of analysis for the CMA. Past studies have reported a wide range of potential breath group lengths in spontaneous speech. For example, Winkworth, Davis, Adams, & Ellis (1995) reported a mean length of 3.84 seconds across six subjects, with a range of 0.3-12.6 seconds for speech at a typical conversational level. How these values scale in a spontaneous speech task with explicit changes in vocal effort has to our knowledge only been investigated in the related context of the Lombard effect (Winkworth & Davis 1997), but the researchers did not find significant changes in breath group length or inspiratory lung volume as a function of speech intensity as has been reported in other studies (Hixon, Goldman, & Mead 1973). However, the functional communicative demands, and therefore the physiological resources mustered for the production of speech-in-noise, could differ from those of projection of speech at a distance. It's possible that both inspiratory lung volume and breath group length may have changed significantly had even higher levels of vocal effort/speech intensity been required of the speakers in the Winkworth & Davis (1997) study.

Given the lack of a clear understanding of the relationship between vocal effort and breath cycle length (or inversely, respiratory rate), and our lack of a direct measure of respiration, an estimation of average breath-cycle period length in each effort condition was determined from the video recordings for one participant in order to determine the proper offset range for use in the correlation map analysis. Using ELAN software, inspirations were coded for each trial. The number of inspirations for a trial was then divided by the length of the trial in seconds, yielding an average respiratory rate in Hz. The average breath cycle period T was then determined according to T = 1/f. Averaging across trials within each condition then gives the average breath-cycle period for each effort condition. The nearest whole-second integer to the longest of these average periods was then selected as the offset range for the CMA, in order to ensure that changes in coordination due to modulation of the breath cycle were captured by the analysis.

Results of this analysis are summarized in Figure 3.13. Average breath cycle period decreased with increased vocal effort. This decrease was most pronounced between the 'shout' and 'yell' conditions, with 'yell' showing an average period length of approximately 3 seconds. Eight seconds, the nearest integer to the mean of the 'normal' condition, was selected as the window of analysis for the CMA.



**Figure 3.13.** Average breath cycle period, participant one. Means by vocal effort condition. Error bars are standard deviation. The total number of inspirations for each trial was divided by trial length (seconds) in order to obtain a measurement of average respiratory rate (Hz). Average period was then calculated according to T = 1/f.

Although results are presented for only one talker, the figure is suggestive in that shortening of the breath cycle takes place at increased levels of vocal effort, as originally hypothesized by Winkworth & Davis (1997). It is likely that this is true of all talkers, given the physiological effects of increased vocal effort (Sharp et al. 1975). This is especially evident in the yell condition. Because unscripted spontaneous speech was used, the relatively small standard deviations for the 'normal'-'shout' conditions were somewhat unexpected. At least two possibilities could account for this: 1) organization of the breath cycle over relatively long instances of conversational speech settles into a relatively stable average mode of behavior, despite inevitable fluctuations in inter-breath interval length (Suki 2002; Suki et al. 2003). 2) Because the measurement technique was limited to audio-visual inspection, only those breaths that were large enough to be either audible or visually detectable were counted. The breath cycle may have been

undersampled in the sense that smaller, shorter inspirations were undercounted. For the present analysis, this is not an issue, since undercounting would result in a *longer* average breath cycle period than that which a more accurate measurement would provide. Therefore, the selected window length of 8 seconds for CMA is conservative, and ensures that the period of the breath cycle is fully represented in all effort conditions.

## 3.4 Collective behavior of the Head, Torso, and Speech Acoustics

The results of the analyses outlined in Section 3.1 are presented in the following sections. In Sections 3.4.1 and 3.4.2, the results of the DFA, CMA, and SVD of the correlation map are discussed with respect to the collective action of the head, torso, and speech acoustics, with an aim toward evaluating whether or not mutual entrainment occurs as a result of increasing vocal effort. Section 3.4.3 shows the fluctuation analysis results for the three center-of-pressure measurements, and presents the results of the CMA and SVD assessing the interaction of the acoustic and kinematic measurements with the global center of pressure. Section 3.5 is devoted to the assessment of postural instability, as indexed by the coordination between the individual feet. In addition to the CMA and SVD for the interaction between the center-of-pressure measurements of the results of the PCA for the 6d force plate signals, and DFA results for the individual force and torque signals for each plate.

# 3.4.1 Fluctuation Analysis of Speech and Affiliated Motion Signals

In physiological systems, understanding of a system's fluctuations provides important insight into the processes that generate the measured behavior. Importantly, fluctuations and their characteristic exponents are sensitive to physiological changes. As discussed previously, vocal effort is known to induce several interrelated physiological changes, both postural (Giovanni et al. 2008), and respiratory (Sharp et al. 1975). It stands to reason, then, that changes in the fluctuations of head motion, torso motion, and speech acoustics will systematically vary with different levels of vocal effort.

## 3.4.1.1 General characteristics of the scaling exponent in each measurement domain

Figure 3.14 shows the DFA results for the acoustic measures, head motion, and torso motion. Interestingly, the fluctuations of both F0 and aRMS exhibit long range correlations, with a scaling exponent in the power-law range, with  $0.5 < \alpha < 1$  in all effort conditions. The presence of power-law distributed fluctuations is indicative of self-similar structure, and is consistent with the findings of Voss and Clark (1975), who reported that fluctuations in amplitude for spontaneous speech follow a 1/f distribution. Interestingly, torso motion fluctuations in all effort conditions evidenced a mean scaling exponent very close to  $\alpha \approx 1$ , or 1/f pink noise. This has two important implications. First, spectral power of torso motion is distributed across a wide range of frequency bandiwdths, and second, the statistical properties of this distribution of spectral power are similar across a wide range of scales, as shown in Figure 3.15. 1/f systems lack characteristic, or dominant timescales, which improves their adaptability to perturbations (West & Schlesinger 1990). Thus, the presence of statistical self-similarity in the movement characteristics of the torso suggests greater adaptability to the influence of perturbations associated with speech at high levels of vocal effort.

This contrasts with the scaling exponents calculated for the head, which were in the range  $1 < \alpha < 1.5$ . This indicates that head motion in all effort conditions is dominated by a characteristic long-period (low-frequency) timescale, with high-frequency fluctuations generally playing a less important role in its dynamics in the context of spontaneous speech production. This is consistent with findings reported by Pozzo, Berthoz, & Lefort (1990), who found that power spectra of head motion were dominated by low frequencies in the range from 0.4-3.5 Hz.



**Figure 3.14. DFA results, head, torso, and speech acoustics.** Results by speaking condition. Individual participant results (1-6), and the mean across all participants (black line) for each effort condition.

## 3.4.1.2 Trends and differences across effort conditions

Further inspection of Figure 3.14 shows that across these measurement domains, there is a general trend for the mean value of the scaling exponent to increase at high levels of vocal effort, indicating an increase in signal smoothness and increased concentration of spectral power at low-frequencies. This increase was marginal for aRMS, and within-subjects ANOVA found no effect of effort condition. A main effect of effort condition was found for F0, F(4,19) = 10.2, p < 0.0005,  $\omega_p^2 = 0.56$ , with post-hoc HSD finding 'shout' and 'yell' to both differ from the lowest three effort conditions. An effect of effort condition was also found for head motion, F(4,19) = 4.4, p < 0.05,  $\omega_p^2 = 0.32$ , as well as torso motion, F(4,19) = 3.1, p < 0.05,  $\omega_p^2 = 0.22$ .



**Figure 3.15.** Power law structure and self-similarity of a torso motion signal. Left: fluctuations in the magnitude of torso motion at multiple time-scales exhibit similar structural properties, as shown by successive magnifications of the signal. Right: linear fit of  $F(L) \sim L$  determines a scaling exponent  $\alpha = .9487$ , indicating power law structure of fluctuations. Bottom: the self-similarity of the signal is evident in the similarity of the histograms calculated for different temporal scales, with the parameter  $\lambda$  denoting the number of subintervals, and the frequency counts of the histograms scaling as  $\lambda^{-\alpha}$ 

The trend for values of the scaling exponent to increase at the two highest levels of vocal effort across both acoustic and kinematic measurement domains has several implications. First, signal fluctuations become smoother at high levels of vocal effort due to the increased influence of longtimescale fluctuations relative to short-timescale fluctuations. Moreover, the commonality of the pattern across these different domains suggests a mutual influence; namely, increased strength of the lowfrequency driving force of the respiratory system associated with the modulation of vocal effort. However, the presence of this trend does not in and of itself guarantee that fluctuations across these systems cohere in such a way as to result in more coordinated behavior with increased vocal effort.

# 3.4.2 Mutual Organization of Fluctuations: CMA

In many physiological systems, the organization of fluctuations in the action of coupled components has important consequences for behavior. For example, in the function of the brain, fluctuations in the period of time between successive firings<sup>2</sup> of individual neurons are known to be due to fluctuations in both membrane potential rise, and fluctuations in the threshold level that the membrane potential must achieve for neural discharge to take place (Musha & Yamamoto 1997). Moreover, in pairs of neurons in visual cortex, membrane potential fluctuations are known to be strongly correlated, and the strength of these correlations increases in the presence of visual stimulus (Lampl et al. 1999). Although speech production is qualitatively quite different from vision, both neuro-physiological processes share a commonality in that each requires the mutually coordinated action of the behaviors of many interconnected components, with general principles of spatio-temporal organization underlying the action of both systems.

In this sense, characterization of the fluctuations of the individual components of a system, and more importantly, of the mutual patterning of these fluctuations with respect to one another, provides a general means of understanding the underlying dynamics of coordination across a wide variety of living systems. In the present study, correlation map analysis (CMA) was used to determine whether or not the fluctuations of the head, torso, and speech acoustics exhibit changes in their mutual organization as a function of vocal effort.

As can be seen in Figure 3.16, CMA shows that the common 'smoothing' trend in the fluctuations of the individual signals is also associated with tighter spatio-temporal coordination across measurement domains. The figure contains individual plots for the m%COR-pos and m%COR-neg measurements, as well as a separate plot for the results of the SVD. All plots display effort condition means. Error bars have been omitted in order to aid visual clarity. Because the pictured means in Figure 3.16 are calculated across speakers, there is an inevitable degree of variability within each effort condition<sup>3</sup>. However,

<sup>&</sup>lt;sup>2</sup> Like speech, inter-spike interval fluctuations follow a 1/f power-law distribution (Musha 1997). The appearance of 1/f distributions is often associated with sufficiently complex structures or processes involving many interlinked components operating at many different characteristic spatio-temporal scales. (West & Shlesinger 1990)

<sup>&</sup>lt;sup>3</sup> Recall that within-subjects ANOVA was conducted to address this source of variation

although within-condition variability is important from the standpoint of entrainment in the sense that we might expect to see decreased variation in these measures for individual speakers in the higher effort conditions, this is a secondary concern to the present analysis. Instead, the visualization highlights an aspect of the data more important to the present analysis of coordination, this being the changes across these pairwise mappings *relative to one another* with increased vocal effort. We predicted that the correspondence would strengthen and simplify across many of these pairwise mappings as a result of entrainment, so it is important to highlight this change-in-concert visually as it is the primary emphasis of the analysis.

If increases in vocal effort result in stronger coordination across the timescale defined by the breath cycle (in other words, the offset range of the CMA), the m%COR measurements are expected to increase, yielding a predicted upward trend with increasing levels of vocal effort. Likewise, simplification of the pattern of coordination will be reflected in a downward trend in the SVD, as fewer singular values will be needed to represent the variance of the correlation map.

In line with our predictions, there is a general tendency across signal pairs for both m%COR-pos and m%COR-neg to increase at the two highest levels of vocal effort. Likewise, the SVD analysis shows a general tendency for systematic simplification in the structure of correlations with successive increases in vocal effort up to the 'shout' level. Interestingly, in the 'yell' condition, only the acoustic-torso pairings showed a continuation of this trend. Signal pairs involving the head (RMS-Head, F0-Head, Torso-Head) or that have a known functional coupling to the head (F0-RMS) notably exhibit more complex coordination relative to 'shout', although this coordination is still simpler than that found in the lowest three effort conditions. An overview of statistical results for the individual signal pairs follows. For a complete summary, the reader is referred to the tables in Appendix A.



**Figure 3.16. Acoustic-Kinematic CMA and SVD results**. Positive correlations, negative correlations, and SVD as a function of effort condition. Markers indicate means across all participants. The strength of positive and negative correlations is indexed by the m%COR measurement, as described in section 3.1. The total number of possible singular values is 961, which is the number of offsets in the correlation map, and corresponds to the analysis window length: +/- 8 seconds.

#### 3.4.2.1 FO-RMS

For the two acoustic measures, no significant difference in the strength of positive correlation, m%CORpos, was found across effort conditions. However, within-subjects ANOVA showed an effect of vocal effort condition on the strength of negative correlations, F(4,79) = 8.9, p < 0.0001,  $\omega_p^2 = 0.26$ . Post-hoc HSD determined m%COR-pos in 'yell' to be significantly greater than 'normal', 'loud', and 'shout'. For the SVD of the correlation map, a significant effect of effort condition was found, F(4,79) = 6.1, p < 0.0005,  $\omega_p^2 = 0.19$ . In this case, Tukey's HSD showed that the number of singular values in 'shout' was significantly less than in 'normal' and 'loud'.

## 3.4.2.2 Speech Acoustics-Head

For F0-Head, effort condition was significant for the strength of both positive correlation, F(4,79) = 22.15, p < 0.0001,  $\omega_p^2 = 0.49$ , and negative correlation, F(4,79) = 23, p < 0.0001,  $\omega_p^2 = 0.5$ . In both cases, post-hoc tests showed 'shout' to differ from all other effort conditions. Similarly, the simplification of correlation structure evident in in the SVD results of Figure 3.13 was significant, F(4,79) = 21, p < 0.0001, $\omega_p^2 = 0.47$ . Post hoc HSD showed 'shout' to require fewer singular values relative to all other effort conditions. In addition, 'loud' and 'louder' were also found to differ from one another.

The RMS-Head correlations did not show the same pattern of results as F0-Head. Neither m%COR-pos, nor m%COR-neg differed significantly across effort conditions. However, the SVD results were similar to those for F0-Head, F(4,79) = 9.1, p < 0.0001,  $\omega_p^2 = 0.27$ . Post hoc HSD found similar results to those for F0-Head, with 'shout' requiring significantly fewer singular values than the lowest three effort conditions (normal, loud, louder). A significant difference was also found between 'loud' and 'louder', with 'louder' requiring fewer singular values.

### 3.4.2.3 Speech Acoustics-Torso

F0-Torso showed a significant effect of effort condition for the strength of positive correlation, F(4,79) = 11.63, p < 0.0001,  $\omega_p^2 = 0.32$ . Post-hoc HSD found m%COR-pos in both 'shout' and 'yell' to be

significantly higher than the lowest three effort conditions. Similarly, the strength of negative correlation was significantly different across effort conditions, F(4,79) = 14, p < 0.0001,  $\omega_p^2 = 0.37$ . In this case, m%COR-neg in 'Yell' was significantly higher than all other effort conditions, while 'shout' differed from 'normal' and 'loud'. For the SVD, there was a significant effect of effort condition, F(4,79) = 7.1, p < 0.0001,  $\omega_p^2 = 0.21$ , with significantly fewer singular values necessary in 'yell' relative to 'normal' and 'louder', while 'shout' required significantly fewer than 'normal'.

RMS-Torso showed similar effects. The strength of positive correlation was significantly different, F(4,79) = 7, p < 0.0001,  $\omega_p^2 = 0.21$ , with 'yell' differing from all other conditions. There was also an effect of effort condition for negative correlation, F(4,79) = 5.3, p < 0.001,  $\omega_p^2 = 0.16$ . As with F0-Torso, m%COR-pos was significantly higher in 'yell' relative to all other effort conditions. For the SVD, the effect of effort condition was significant, F(4,79) = 6.6, p < 0.0005,  $\omega_p^2 = 0.2$ , with the loudest three effort conditions requiring fewer singular values than 'normal'.

Of all signal pairings, the gross difference across effort conditions in the number of singular values necessary to reconstruct the data with 99.8% accuracy was greatest in the acoustic-torso pairs. Visually observable body motion is intermittent in its coordination with speech at the 'normal' level of vocal effort. But these results suggest that this intermittent action becomes increasingly coactive and coordinated with the acoustic measures at high levels of vocal effort.

### 3.4.2.4 Torso-Head

Positive correlation increased significantly at higher levels of vocal effort, F(4,79) = 6.6, p < 0.001,  $\omega_p^2 = 0.2$ , with 'shout' and 'yell' showing significantly higher values of m%COR-pos than the lowest three effort conditions. A nearly identical effect was found for negative correlation, F(4,79) = 6.7, p < 0.0005,  $\omega_p^2 = 0.2$ , again with 'shout' and 'yell' differing from the lowest three effort conditions. SVD showed a trend toward simplification up to 'shout', with a significant effect of effort condition F(4,79) = 6.8, p < 0.0005,  $\omega_p^2 = 0.2$ , again with 'shout' and 'yell' differing from the lowest three effort conditions.

0.0001,  $\omega_p^2 = 0.21$ . Post-hoc HSD found 'shout' to differ from both 'normal' and 'loud', while 'normal' differed from 'louder'.

## 3.4.2.5 General Remarks

The combined results of the CMA and DFA support the hypothesis that increased vocal effort results in tighter spatiotemporal coordination between head motion, torso motion, and speech acoustics. However, the analysis uncovered several interesting points of departure from this general picture. Combining the observations outlined above, it is clear that the system-wide correlation structure shows a trend toward simplification up to the 'shout' level, as this tendency was evident across all of the discussed signal pairs. But interestingly, this tendency for simplification continued all the way through 'yell' only for the acoustic-torso pairs.

This raises two points of particular interest: 1): the nature of acoustic-motion coupling appears to undergo an unexpected shift in the transition from 'shout' to 'yell', becoming less correlated with the motion of the head while becoming more correlated with the motion of the torso. And 2): In 'yell', the F0-RMS pairing showed a significant increase in the strength of *negative correlation*. This effect was paralleled in the acoustic-Torso pairs. Notably, the strength of negative correlation only increased in 'yell' for RMS-Torso. However, this effect was noticeably absent in the acoustic-Head pairs, despite being present in Torso-Head. This suggests that the evident simplification in the coordination between the torso and speech acoustics in 'yell' is the effect of an underlying symmetric periodicity driving their mutual action, which is likely due to the influence of the respiratory perturbation due to vocal forcing. Interestingly, this same effect is not reflected in the behavior of the head with respect to the organization of the speech signal.

# **3.5**Coordination with global COP

The results of the DFA for the center-of-pressure time-series are shown in Figure 3.17. The value of alpha for each of the individual plate COPs and the global COP trended higher in 'shout' relative to all

other effort conditions, including a marginal increase relative to 'normal' for COP one, but withinsubjects ANOVA found no effect of effort condition for any of the three signal types.



**Figure 3.17. DFA Results, COP time series.** Values of the scaling parameter  $\alpha$  by effort condition. Results for participants 1-6, plus the mean across all participants (black line).

In spite of this finding, the global COP measurement evidenced significant correlation with motion of the head and torso in the CMA. These results are shown in Figure 3.18.

For all effort conditions, COP-Head showed the highest overall level of correlation among all signal pairs. This is indicative of the importance of the head for postural control. This was true for both positive correlation and negative correlation. Both positive and negative correlation display a trend similar to that found for the acoustic-head pairs discussed in the previous section, with the strength of correlation increasing up through the 'shout' level, and then decreasing for 'yell'. The differences in

effort condition means were significant for both positive correlation, F(4,79) = 4.37, p < 0.01,  $\omega_p^2 = 0.13$ , and negative correlation, F(4,79) = 3, p < 0.05,  $\omega_p^2 = 0.083$ . These effects were mirrored by the SVD, which showed significant reduction in the number of singular value components up through 'shout', followed by a subsequent rise in the number of components in 'yell', F(4,79) = 11, p < 0.0001,  $\omega_p^2 = 0.3$ .

COP-Torso showed trends similar to the acoustic-torso pairs for the strength of positive and negative correlation, with an increase in the two highest effort conditions. Within-subjects ANOVA determined these effects to be significant for positive correlation, F(4,79) = 6.67, p < 0.001,  $\omega_p^2 = 0.2$ , and for negative correlation, F(4,79) = 7, p < 0.0001,  $\omega_p^2 = 0.21$ . For both types of correlation, post-hoc HSD differentiated 'shout' and 'yell' from the lowest three effort conditions. While SVD showed simplification up through 'shout' with a significant effect of effort condition F(4,79) = 4.8, p < 0.005,  $\omega_p^2 = 0.14$ , this simplification did not continue through the 'yell' condition.

On the surface, this observation of stronger correlation paired with increased complexity might appear contradictory. However, a similar effect was observed for the Torso-Head pair in the previous section. Furthermore, the strength of correlation between head motion and the COP is two to three times stronger than that between torso motion and the COP. Despite the increased strength of correlation between the torso and the COP at high levels of vocal effort, the head continues to exert a stronger influence on the COP time series. Indeed, in this study, the complexity of correlation structure is generally lower in signal pairs involving the head, as compared to the other signal pairings. This underscores the importance of the head in both vocalization and postural control, and its important role as a pivot-point in mediating the crossover of behavior between these task domains.

Regarding coordination of the COP with the acoustic measures, F0-COP showed a pattern on par with F0-Head, with the strength of positive and negative correlation increasing up through 'shout'. Significant effects of effort condition were found for both positive and negative correlation, F(4,79) = 8.9, p < 0.0001,  $\omega_p^2 = 0.26$ , and F(4,79) = 7.6, p < 0.0001,  $\omega_p^2 = 0.23$ , respectively. Post hoc HSD showed the m%COR measures to be significantly higher in 'shout' relative to the lowest three effort conditions. For

the SVD, the trend from lower to higher vocal effort shows a simplification of correlation structure, followed by a large increase in complexity for 'yell'. Once again, this effect was significant, F(4,79) = 13, p < 0.0001,  $\omega_p^2 = 0.36$ . Post hoc HSD showed that the correlation structure for 'yell' was significantly more complex than for 'loud', 'louder', and 'shout', while 'shout' was significantly simpler than 'normal' and 'loud'.

Conversely, for RMS-COP, the strength of correlation decreased as vocal effort increased. Also, the effect of effort condition was relatively weak for positive correlation, F(4,79) = 2.8, p < 0.05,  $\omega_p^2 = 0.075$ , and negative correlation: F(4,79) = 2.7, p < 0.05,  $\omega_p^2 = 0.072$ . The relatively weak correspondence between RMS-COP parallels the relatively weak correspondence between RMS-Head, and could reflect the dual involvement of the head in vocalization and in postural control. That is, the head is likely a mediating link in the coordination observed between speech acoustics and global COP across different vocal effort levels.



Figure 3.18. Center of pressure CMA and SVD results. Positive correlation, negative correlation, and SVD as a function of effort condition. Markers indicate means for each condition across all participants.

## 3.6 Postural instability

The corollary hypothesis predicted that raising vocal effort past a critical threshold would induce high levels of within-speaker coordination, and that this coordination would ultimately impose a restriction on the ability of the body to maintain balance. This effect would be evidenced by diminished coordination in the bodily subsystems responsible for coordination at the point of environmental contact, that is, between the individual feet. This cross-domain correspondence would likely be accompanied by diminished within-domain correspondence in each of the individual feet as well.

Figure 3.17 showed that the patterning of fluctuations for the individual force plates changed little with increased vocal effort. Although no main effect of effort condition was found for the individual plate COP fluctuations, the nature of the coordination between the feet did in fact change. Figure 3.19 shows that the strength of positive and negative correlation between the center of pressure measurements for the individual feet tended to increase with increased vocal effort up to 'shout', but substantially decreased in 'yell'. These effects were significant: positive correlation: F(4,79) = 6.41, p < 0.001,  $\omega_p^2 = 0.2$ , negative correlation: F(4,79) = 5, p < 0.005,  $\omega_p^2 = 0.15$ . Post-hoc tests showed that positive correlation was significantly stronger in 'louder' and 'shout' than in the other three effort conditions. For negative correlation, 'louder' and 'shout' were different from 'normal' and 'yell'. The SVD results show a parallel effect, with simplification of the correlation map up through 'shout', and a subsequent increase in complexity in 'yell', F(4,79) = 7.1, p < 0.0001,  $\omega_p^2 = 0.22$ .



**Figure 3.19. CMA and SVD results, COP1-COP2.** Left: m%COR-pos and m%COR-neg results, error bars are standard deviation. The strength of correlation significantly decreases for both measures in 'yell'. Right: cumulative variance plot for SVD of COP1-COP2. More singular values are required to reconstruct the data in 'yell' with 99.8% accuracy relative to all effort conditions except 'normal'. However, note that at even higher levels of accuracy (eg. 99.95%), this also becomes true relative to 'normal'.

# 3.6.1 Principal component analysis of 6d force plate measures

Given the apparent conflict between the results of the fluctuation analysis for the individual plate COP measurements and the results of the CMA, PCA was performed on the original 6d force plate signals in order to gain a better understanding of the reduction of coordination between COP1 and COP2 in the 'yell' condition. The force plate measurements amount to transductions of the collective action of numerous neuromuscular components involved in postural control. This suggests that a degree of structural similarity should be present among the components of the 6d signals, that is, among the individual forces (Fx, Fy, Fz) and torques (Mx, My, Mz). This means that each 6d signal can in principle be represented in terms of a lower dimensional combination of variables that correspond to the underlying processes responsible for this redundancy.

Table 3.3 contains the mean number of principal components necessary to explain 96% of the variance in each of the 6d force plate signals. For both plates, more principal components are necessary on average to represent 96% of the variance in the original data in 'yell' than in the other effort conditions. This effect is more pronounced for FPL-two, but within-subjects ANOVA showed significant effects of effort condition for each foot. For plate one, F(4,79) = 3.4, p < 0.05,  $\omega_p^2 = 0.1$ , and for plate

two, F(4,79) = 5.3, p < 0.001,  $\omega_p^2 = 0.16$ . This is suggestive of diminished system-internal correspondence among the individual channels of the 6d force plate signals for both force plates.

Analysis of the eigenvalue spectrum showed that the summary view of this effect presented in Table 3.3 also applies at the level of the individual components. Figure 3.17 contains cumulative variance plots for each force plate. For FPL-one, less cumulative variance is explained by the addition of each principal component in 'yell', with the exception of reconstructions consisting of five components. For reconstructions consisting of two, three, and four components, 'yell' differed significantly from other effort conditions<sup>4</sup>, though exactly which conditions showed significant differences from 'yell' was dependent upon the number of principal components. This effect was strongest in the four component reconstruction, F(4,79) = 6.1, p < 0.0005,  $\omega_p^2 = 0.19$ , which corresponds approximately to the 96% variance threshold in all effort conditions. In this case, post-hoc HSD showed that 'yell' differed from 'loud', 'louder', and 'shout.

**Table 3.3 Force plate PCA results.** Mean number of principal components necessary to explain 96% of the variance in the 6d forceplate signals for each effort condition. More principal components are necessary to capture 96% of the variance of the 6d force plate signals in 'yell' relative to the other effort conditions.

	Condition	Mean # PCs
FPL-One	Normal	4.4
	Loud	4.1
	Louder	4
	Shout	4.2
	Yell	4.6
FPL-Two	Normal	4.4
	Loud	4.2
	Louder	4
	Shout	4.3
	Yell	5

<sup>&</sup>lt;sup>4</sup> Full results appear in appendix C, including cumulative variance plots for individual participants.
For FPL-two, this tendency was true of reconstructions consisting of two or more components. Furthermore, 'yell' differed significantly from *all* other effort conditions in each of these cases. Again, this effect was strongest for the four component reconstruction, F(4,79) = 11, p < 0.0001,  $\omega_p^2 = 0.3$ .

As a final step, analysis of the eigenvectors of the principal components was performed in order to identify the variables in the original force plate signal that contributed most to the reduced representation of the data produced by PCA. Figure 3.20 shows these results. For each of the force and torque measurements in the original 6d signals, the mean of the squared loading was computed for each effort condition. Recall that the squared loading indexes the strength of the individual variable's contribution to a given principal component. With respect to the figures, the values shown are averaged across all principal components used in the reduced representation of the data accounting for 96% of the variance. That is, the number of principal components necessary to explain 96% of the variance for a given trial is first determined. Then, the loadings for each of the principal components are squared. These values are then averaged, yielding the average squared loading for each variable for the trial. Finally, the results of the individual trials are then grouped by effort condition, and averaged.



**Figure 3.20. PCA Cumulative variance, force plates one and two.** Mean variance explained by each principal component. Less variance is explained by each principal component in 'Yell' relative to the other effort conditions.

Inspection of Figure 3.21 shows that for both force plates, the three component force signals (Fx, Fy, and Fz) tend to contribute most to the principal components in the three lowest effort conditions. In the higher vocal effort conditions, the contribution of Fz is supplanted by Mx in 'shout', and Mz in 'yell'. Furthermore, in 'yell', the contribution of Mx is greater than or equal to the contribution of Fz in force plate one and force plate two, respectively. This points to a functional reorganization of the postural control system, which adapts to the task demands of speech at high levels of vocal effort, and is suggestive of a reconfiguration of neuromuscular recruitment needed for postural control: a change in synergistic response indexed by change in ground reaction forces and moments.



**Figure 3.21. Average squared loadings of force plate variables.** Means by effort condition, force plate forces (Fx,Fy,Fz) and torques (Mx, My, Mz) for each force plate.

#### 3.6.2 DFA results, Individual Forces and Torques

Reduced correspondence between the channels of the 6d force plate signals points to a shift in their fluctuation patterns. DFA of these individual force and torque signals corroborates this notion. These results are displayed in Figure 3.22. With the exception of the forces in the y direction for plates one (Fyone) and two (Fy-two), all force plate signals exhibited a scaling exponent greater than 1 in all effort conditions. For Fy-one, alpha was greater than 1 in 'shout' and approximately equal to 1 in all other conditions. For Fy-two, alpha was also greater than 1 in 'shout', and less than 1 in all other conditions. All measures showed a significant effect of effort condition, with the exception of My and Mz for force plate one, and Mx for force plate two. In most variables, post-hoc tests determined 'shout' to differ

significantly from 'yell', though many other differences exist for several variables. The reader is referred to the table contained in Appendix B for a full listing of ANOVA results.



**Figure 3.22 DFA results, force plate component forces and torques**. Each plot contains effort condition means. The table represents the mean value of the scaling exponent (SD) averaged across all signals.

There is a general tendency for 'shout' to exhibit the highest value of alpha for all signals, with the exception of Mz-two. In conjunction, the majority of signals display a subsequent drop off in 'yell', indicating rougher signals with a distribution of power across a wider spectrum of frequencies.

Interpreted in conjunction with the results of the CMA and the PCA, these results show that the reduced coordination between COP1 and COP2 in 'yell' is symptomatic of changes in the characteristics of the fluctuations of the individual components. Much like the mutual smoothing of fluctuations among torso motion, head motion, and speech acoustics led to improved coordination, mutual smoothing among the components of the 6d force plate signals is associated with improved coordination between the feet in the 'shout' condition. But the improved coordination in 'louder' relative to other effort conditions evident in the CMA demands a slightly different explanation. The PCA results suggest that 'louder' displays better correspondence among the individual variables than the other effort conditions. But the DFA results show no effect of mutual smoothing. However, inspection of the mean and standard deviation of

the scaling exponent across all variables in Figure 3.22 shows that both 'louder' and 'shout' exhibit the smallest spread of the scaling exponent (SD). This suggests that at least two characteristics of fluctuations are associated with improved coordination across an ensemble: i) a mutual increase in the scaling exponent, or in other words mutual smoothing across the collection of signals, and ii) attraction of the scaling exponents to a common value or behavioral regime.

Conversely, in the 'yell' condition, the marked decrease in the value of the scaling exponent across these signals is mirrored by the results of the PCA and CMA/SVD. Furthermore, it is interesting to note that in the case of FPL-two, the increased value of alpha for Mz in 'yell' contrasts sharply with the trends evident in the other variables. For FPL-two, recall that more principal components were necessary to explain 96% of the variance of the original data, and that less variance was explained by each additional component relative to FPL-one (see Table 3.3 and Figure 3.20), pointing to less structural coherence among the individual variables. The fact that the PCA eigenvector analysis shows Mz to contribute more to the underlying dynamics in this condition in FPL-two suggests that tension between the fluctuation patterns of the individual components could be an additional cause of this effect.

#### **Chapter 4: Discussion**

As hypothesized, the present work shows that increasing vocal effort during speech production entrains multiple bodily subsystems. Specifically, at high levels of vocal effort, kinematic measures of body and head motion become more tightly coordinated with each other and with acoustic measures of the voice. Increased vocal effort resulted in increased coordination in the correlation map analysis, as indexed by the m%COR measurement for both in-phase and anti-phase coordination, and reduced the number of factors accounting for the increased correlation, as shown by singular value decomposition. This indicates that the time-varying correspondence between these measurement domains both strengthened and simplified with increased vocal effort. Observed changes in the coherence between measurement domains, as shown by the results of the detrended fluctuations within the individual measurement domains, as shown by the results of the detrended fluctuation analysis (DFA). Increasing vocal effort resulted in increased values of the scaling parameter alpha, indicating smoother signal profiles due to the increased influence of long-timescale fluctuations relative to short-timescale fluctuations. The fact that this effect was both general across the several subsystems and paired with improved correspondence between the subsystems suggests a mutual perturbing influence due to increased vocal effort.

The results of the analysis also supported the corollary hypothesis that within-speaker coordination would ultimately conflict with environmental coordination at the highest levels of vocal effort, resulting in postural instability. In the 'yell' condition, this instability was evident in the decreased strength of coordination between the center of pressure measurements for the individual feet, and increased complexity of correlation structure. Furthermore, PCA decomposition of the variance for the 6d force plate signals for each foot in this condition showed evidence of poorer correspondence between the forces and torques. More principal components were required to represent 96% of the variance of the 6d signals in 'yell' relative to the other effort conditions, and comparatively less variance was explained by each principal component. Analysis of the PCA eigenvectors uncovered a notable difference in the contribution of the six force and torque components to the PCA representation in the two highest effort

conditions. In 'shout', Mx, which indexes the postural rotation around the horizontal axis, or anteriorposterior sway, contributes more variance to the PCA representation. In 'yell', the z-torque Mz, or the rotation of the ankles around the vertical axis, contributes more variance.

Interestingly, DFA of the individual channels of the 6d force plate signals showed decreased values of alpha in 'yell' relative to 'shout' for all channels in both force plate signals except for the z-torque of plate two. The decrease of scaling exponents as vocal effort increased indicates an increase in the roughness of the signals in these individual channels, with spectral power distributed across a wider range of frequencies. The divergence of the z-torque for plate two from this general pattern, combined with its increased contribution to the variance in this effort condition, is likely at the heart of the weaker PCA results for force plate two as compared to plate one. This contrasts with the effects seen in the body-head-acoustic comparisons, where similar changes in fluctuation patterns were associated with stronger cross-domain correspondence. In other words, similarity in the observed changes of fluctuation patterns of the subcomponents of the scaling exponent among subcomponents of the foot-foot postural subsystem are instead associated with poorer time-varying correspondence.

In sum, the dual hypothesis that increased vocal effort would result in mutual entrainment among head motion, torso motion, and speech acoustics, and ultimately a concomitant loss of postural stability, was well supported by the results of these analyses. However, the initial hypothesis that each successive increase in vocal effort level would be accompanied by increased correspondence between measurement domains turned out to be overly simplistic. Several points in this regard merit further discussion. First, there is the question of the large increase in strength of coordination in 'shout' relative to the smaller increases seen in the changes between the lower three effort conditions. Second, there is the question of the difference between 'shout' and 'yell', where coordination between some subsystems continues to strengthen and simplify, while in others, this is not the case. Finally, the issue of postural instability merits discussion in terms of systems-level processes that explicitly address the apparent conflict between high levels of intra-speaker coordination and coordination of the talker with the environment.

#### 4.1 Increased coordination in 'shout' relative to the lowest three effort conditions

Although the results showed a trend toward stronger coordination characterized by lower dimensional relations and corresponding changes in fluctuation patterns between measurement domains at successively increasing levels of vocal effort, this effect was especially pronounced in the 'shout' condition. This raises the question of whether the observed changes in coordination are linked to vocal effort via some kind of gradient, and to what extent such a gradient may or not scale linearly as vocal effort increases.

As a first step toward addressing this issue, it is important to reiterate that the elicitation of vocal effort in this study involved varying the intended communicative distance associated with speech production (Traunmüller & Eriksson 2000). In the present experiment, distances of 1.5m, 4m, 7.5m, and 30m were used. Where the lowest three effort levels involved approximately 2 to 2.5-fold increases, the increase from 'louder' to 'shout' involved a 4-fold increase in distance. This large magnitude increase in distance was paired with a similarly large increase in the average acoustic RMS values shown in Figure 3.11. The mapping between acoustic RMS values and communication distance appears consistent, and supports, or at least does not contradict, the possibility that vocal effort indexes a speech parameter that varies along a continuous gradient, as opposed to a system of categorically discriminated settings. However, it should also be noted that increases in vocal effort may scale non-linearly (e.g. possibly according to a power law) with increasing communication distance, and that this scaling is also evident in the changes in coordination observed in the present study. This interpretation would be consistent with the fact that the intensity of sound diminishes with respect to distance from its source according to the inverse square law, a power law with scaling exponent 2, given by Equation 4.1,

$$I \propto \frac{1}{r^2},\tag{4.1}$$

where I is the intensity of sound at a point that is a distance r from the source. This would account for the pronounced difference between shout and the lowest three effort conditions, while still affording a gradient interpretation of the phenomenon.

It is also possible that increasing vocal effort arouses activity in the CNS, or modulates what Lashley (1951) originally referred to as the *dynamic level*, of neuronal activation. For Lashley, who was one of the first to propose that the nervous system functions as a constantly active, integrated network, modulation of the dynamic level of neural activation was crucial to modifying what he called *patterns of facilitation*, or the patterns of neuronal activity that enable complex temporally integrated behaviors. Lashley identified rhythmic activity as a fundamental instance of this process, where temporally spaced waves of facilitative excitation spread throughout the entire nervous system. As he put it:

Consideration of rhythmic activity and of spatial orientation forces the conclusion, I believe, that there exist in the nervous organization, elaborate systems of interrelated neurons capable of imposing certain types of integration upon a large number of widely spaced effector elements; in the one case transmitting temporally spaced waves of facilitative excitation to all effector elements; in the other imparting a directional polarization to both receptor and effector elements. These systems are in constant action. They form a sort of substratum upon which other activity is built. They contribute to every perception and to every integrated movement. (p. 127-128)

In other words, a rhythmic discharge of sufficient strength could integrate the action of many subsystems into coordinated action at the level of behavior. A perturbation exceeding a critical threshold could easily trigger a wave of excitation that would then subsequently travel throughout the CNS (Bassingwaithe et al. 1994). Although the relation of vocal effort to changes of communicative distance and acoustic loudness may in and of itself be gradient, there could be critical points on the gradient where the level of CNS activation reaches a point that facilitates widespread integration of behaviors across the entire body. On the surface, findings that report increased vocal effort to have a positive effect on speech and postural symptoms associated with Parkinson's disease (Ramig et al. 2001) would seem to support this idea. In any case, clarifying whether or not vocal effort exhibits gradient effects on coordination and whether or

not such critical points in CNS activation due to changes in vocal effort exist is beyond the scope of the present research, but future work could be devoted to clarifying these issues.

# 4.2 Difference between 'shout' and 'yell': Respiratory perturbation and Fluctuation Characteristics of the Head and Torso

Interestingly, this trend toward stronger, simpler system-wide coordination up through the 'shout' condition did not simply continue through 'yell'. Here, a clear difference was observed between the behavior of the head and torso. In 'yell', the strength of coordination between the head and all other measurement domains except the torso decreased, whereas coordination between the torso and the other measurement domains continued to both strengthen and simplify.

The relation between these kinematic measures and speech acoustics is of particular interest. Prior work has demonstrated the existence of a functional coupling between rigid body motion of the head and F0 (Yehia et al. 2002). The current investigation corroborates this finding, and furthermore, demonstrates that the strength of this coupling increases with increased vocal effort. But when vocal effort increases past a threshold level, the scaling of this functional coupling between the head and speech acoustics ceases, and correlation between the two domains subsequently diminishes. Physiologically, this might be attributed to the stiffening of the muscles of the neck that takes place with increased vocal effort, (Pettersen, Bjørkøy, Torp & Westgaard, 2005).

A parallel consideration is the nature of the fluctuations of the head uncovered by the DFA. In all effort conditions, the scaling parameter of head motion was  $1 < \alpha < 1.5$ , and the value of the scaling exponent increased with vocal effort level, suggesting an increased concentration of power at low-frequency bandwidths in the head motion signal. The motion of the head approaches a brown noise spectrum, meaning that fluctuations in the long time scale processes of head motion are highly correlated. However, this also means that the system is less adaptive to the influence of perturbations relative to those in which power is distributed proportionally across all timescales (i.e. 1/f pink noise) (Bassingthwaighte et al.,1994; West & Shlesinger, 1990). A perturbation delivered near the system's natural frequency, or

characteristic time scale, will propagate throughout the system, influencing the global integrity of its organization. The low-frequency perturbation due to respiratory forcing at high vocal effort, then, can be interpreted as a destabilizing force with respect to the head when the strength of this perturbation passes a critical threshold. This affects the functional coupling of the head with speech, given that the ability of the head to adapt its behavior in the face of such perturbations is limited. Instead, the response of the head involves a stiffening of the muscles of the neck due to the demands of postural control (Giovanni 2008), which would account for an inhibition of the functional coupling between the head and speech acoustics.

This contrasts with the expected behavior of the torso given the results of the DFA. In general, fluctuations in torso motion were found to exhibit 1/f power-law structure, with a mean scaling parameter  $\alpha \approx 1$ . This reflects similar results reported by Torre, Delignieres, and Lemoine (2007), who found a similar effect for the motion of the hands in bimanual coordination. In the present case, the effect should follow straightforwardly from the physiological composition of the limbs and torso, which consist of proportional body segments of varying length. In such a system, movements of the smaller elements would contribute small-scale fluctuations, with the scale of fluctuations increasing with increasing segment size and length.<sup>5</sup> The presence of this statistical self-similarity in the structure and movement characteristics of the torso suggests greater adaptability to the influence of perturbations due to respiratory forcing at high levels of vocal effort. This adaptability manifests in the simplification of coordination between torso motion and the other measurement domains, even in the 'yell' condition.

The effects of the head's response to the vocal perturbation in 'yell' were also evident in its coordination with the global COP measure. In this condition, the coordination of the head with global COP diminished both in terms of the strength of coordination (m%COR), and in terms of the simplicity of the correlation map. Similar to the acoustic measurements, coordination of the torso with global COP

<sup>&</sup>lt;sup>5</sup> The study of the proportional distribution of human body segment lengths can be traced back to Da Vinci's 1490 *Vitruvian Man.* An apt metaphor for the distribution of the scale of fluctuations in the movements in a proportionally distributed system like the torso is the adaptive response of a tree to a strong gust of wind. In response to the wind, the leaves vibrate quickly at very small magnitudes. The smaller branches sway back and forth at a lower frequency and greater magnitude than the leaves. As the branch size increases, the magnitude of sway increases and the frequency decreases, ultimately giving rise to a 1/f distribution of fluctuations. Recent work in physics has even suggested that the self-similar geometrical structure of trees (saying nothing of their fluctuations), also studied by da Vinci, is related to their need to adapt to wind related perturbations (Eloy 2011)

strengthened in 'yell'. But interestingly, this effect was not paired with simplification of the correlation map. This is likely due to the fact that the head exerts a stronger influence over the global COP than the torso does, as shown by the values of m%COR and the overall number of singular values used in the SVD representation of the correlation maps for COP-Head. Values of m%COR are much higher for the head than for the torso, and roughly half the number of singular values is needed to capture 99.8% of the variance. In spite of this, like its relation with both F0 and RMS, torso motion scales with global COP up through the 'yell' condition. In contrast, the head does so only to a limited extent. Again, this may be due to the torso's adaptability to a low-frequency perturbation delivered by vocal forcing, as quantified by its fluctuation characteristics. The head, on the other hand, does not adapt as readily. When the strength of the perturbation passes a critical threshold, the inability of the head to readily adapt its behavior contributes to postural instability.

Although the results reported in this study did not include a direct measure of respiration, the combined results of the DFA and CMA suggest that a common influence is causing signal smoothing across a wide variety of physical subsystems and stronger, simpler coordination between these subsystems with increased vocal effort. Increasing vocal effort imposes a unique set of functional demands on the respiratory system during speech production by increasing the speed of inspiration and expiration (Hixon 1973), which physiologically equates to the forced expiration of air from the lungs (Sharp 1975). It is likely that the common influence is a low-frequency perturbation affiliated with this respiratory forcing induced by changes in vocal effort. The increase in the values of the scaling exponents uncovered by DFA suggests that energy from such a low-frequency perturbation is a mutual influence on the speech signal and movement of the head and torso. When the perturbation becomes strong enough, the differences in the fluctuation characteristics of the head and torso emerge as differences in coordination. Although the dependence of these effects on increased vocal effort are robust, obviously a respiratory measure should be included in future versions of this work to confirm that changes in respiration are indeed the common influence driving the observed changes in coordination.

#### 4.3 Steps toward a systems view of postural instability induced by increased vocal effort

Ultimately, the appearance of postural instability at the highest levels of vocal effort is due to a reduction in the requisite variety, or range of system-internal variation, that an organism must have at its disposal in order to coordinate its behavior with the environment (Ashby 1958). With increasing strength of the perturbation due to respiratory forcing, the action of multiple bodily subsystems becomes more tightly coordinated, especially the torso and head (cf. Figure 3.16). This entrainment severely limits the degrees of freedom available to the postural control system, thus restricting the ability of the body to maintain balance, that is, to adapt and coordinate its behavior with the environment. This loss of variety is evident at the system boundary, where coordination between the individual feet, a pre-requisite to the maintenance of balance, diminishes.

This process of entrainment constitutes a form of system-wide integration at the level of the physical body. Incredibly, coordination at this lower physical level is intimately related to the needs of information projection, that is, to the task demands of communication, at a higher, dyadic level of interaction. In this way, coordination at the physical level could be constrained by the magnitude of information transfer (Kolasa & Pickett 1989), that is, by the *energetic* requirements of projecting linguistic information as speech across a given distance. With vocal effort, we have direct evidence that the requirements of communication at the social (i.e. inter-agent) level induce different patterns of coordination at the level of individual physical behavior. Moreover, when the magnitude of information transfer increases past a critical threshold, with increasing communicative distance *and* with a change of discourse style (as in the difference between 'shout' and 'yell'), the physical body begins to separate from its environment. As system-internal physical coordination increases, the system becomes increasingly integrated, and thus, structurally distinct (Kolasa & Pickett 1989) from the environment. This ultimately affects the stability of the organism with respect to its physical surroundings, for the benefit of the creation of communication pathways at the level of social interaction.

#### **5. CONCLUSION**

This study sought to test the dual hypothesis that increasing vocal effort in speech production would result in increasingly rigid coordination of body motions associated with speech, and that this rigid coordination would be accompanied by a loss of postural stability at very high levels of vocal effort. Coordination was assessed with a unique combination of methods for non-stationary time series analysis. This included methods for assessing correspondence across measurement domains (correlation map analysis), dimensionality reduction (singular value decomposition analysis, principal component analysis), and a method for assessing fluctuations within measurement domains (detrended fluctuation analysis).

The results of these analyses provided empirical support for both of the hypotheses. Increasing vocal effort resulted in more rigid coordination, as shown by the combination of increased strength of time-varying correspondence and the evidence of dimensionality reduction. Postural instability was evident in diminished coordination between the individual feet, as well as in diminished correspondence between the component forces and torques of the 6d force plate signals. Coordination showed an important parallelism with the fluctuation properties of the individual subsystems. This was apparent both in terms of the appearance of a system-wide smoothing effect with increased vocal effort, and in terms of the response characteristics of different subsystems. The difference in the adaptive response characteristics of torso and head motion indexed by their fluctuations ultimately had important consequences for coordination when the respiratory perturbation due to increased vocal effort was strong enough.

Although fluctuations have been studied in the coordination literature, (e.g. Schmidt, Carello, & Turvey 1990; Carson, Goodman, Kelso & Elliott 1995), they are often considered simply in terms of fluctuations about some central tendency with respect to a task-related goal in a confined, highly stylized experimental context (see Diniz et al. 2010 for a general review). For example, fluctuations in phase

relations between two limbs when *deliberately* trying to coordinate their movements in an in-phase or out of phase manner, as in Carson et al. (1995).

This study took a different approach, in that we examined fluctuations in the movement characteristics of several *heterogeneous* subsystems in a relatively *unconstrained* experimental context, and how these fluctuations changed with respect to one another in response to the task demand of producing speech at increasingly higher levels of vocal effort. Studying fluctuations across many interrelated, yet heterogeneous, subsystems in this way affords a more nuanced view of the interplay between subsystem fluctuations and coordination in spontaneous human behavior. For example, this study found that fluctuations in torso motion have 1/f structure. This characteristic enabled scaling of the coordination between torso motion and speech, which also has 1/f structure, at the highest level of vocal effort. There is an unquestionable connection between coordination and fluctuations at the physical level, especially in systems composed of more or less identical subunits<sup>6</sup>. But as this study demonstrates, complex human behaviors such as speech involve the coordination of many interlinked heterogeneous subsystems. As such, the analytic methodology outlined in this thesis represents a step forward in the way these concepts can be jointly applied at the level of behavior.

#### **5.1 Future Directions**

Confirmation of the two hypotheses investigated in this thesis highlights the necessity of explicitly incorporating respiration into the investigation of speech motor control. The addition of a direct measure of respiration would enable us to directly test the hypothesis that the effects on within-speaker coordination observed in this study are related primarily to changes in respiratory behavior associated with increasing vocal effort.

In many ways, clarifying the role of the breath cycle in these observed effects is of critical importance. The changes in the organization of the respiratory cycle, including increased speed of

<sup>&</sup>lt;sup>6</sup> The power spectra of collections of neurons in the human brain in subjects at rest has been found to have 1/f structure (e.g. Novikov 1997), and synchronized neural activity facilitates communication among groups of neurons in diverse cortical regions (Varela 2001; Fries 2005).

inspiration and expiration combined with increased respiratory forcing, are thought to be the root cause of the within-speaker entrainment effect, and ultimately, the appearance of postural instability. Once again, the consideration of fluctuations appears to be crucial. Fluctuations in inter-breath interval lengths in healthy human respiration follow 1/f power law structure (Frey et al. 1998). But the periodic breathing exhibited by heart-failure patients, known as Cheyne-Stokes respiration, is associated with a loss of the healthy variability provided by 1/f fluctuations (West, 2013). Cheyne-Stokes breathing does not just involve a loss of variability. Importantly, this loss of variability is due to the development of *large oscillations in the action of the lungs*. This loss then propagates to related subsystems, as it is also evident in the inter-beat-interval fluctuations of the heart beat (e.g. Peng et al. 1995).

We think that an analogous loss of variability in the respiratory rhythm is responsible for the behavioral effects observed in this study. Entrainment to a strong, highly periodic respiratory rhythm would account for the increasingly rigid coordination at higher levels of vocal effort, and the loss of variability associated with entrainment to a specific rhythm would result in postural instability. Thus, the importance of including a direct respiratory measure in future versions of this study cannot be understated. Along these lines, it is also necessary to verify whether or not vocal effort's effects on coordination are truly gradient, or whether critical levels of activation somehow facilitate the high levels of coordination we observed in the 'shout' condition. One way of accomplishing this would be to elicit speech at more distances along the continuum, thereby eliminating the jump from 7.5m to 30m in the present study's transition from 'louder' to 'shout'. This could be accomplished by using an acoustic reverb room in which the acoustic damping could be manipulated to elicit speech at more fine-grained levels of vocal effort.

In this vein, another direction that could be explored is the cross-influence between these physiological considerations and stylistic differences in the aspects of linguistic performance at different levels of vocal effort. For instance, Winkworth et al. (1995) conducted a study of linguistic organization with respect to the organization of respiration for speech produced at a single, comfortable volume, which would equate to the 'normal' level elicited in our study. These researchers found that inspirations tended

to occur at clause boundaries, and that high inspiratory lung volumes resulted in both increased breath group length and longer strings of clausal structures. It's likely that the reorganization of the respiratory rhythm with increased vocal effort has important consequences for linguistic organization (and possibly vice versa). Exploring how the co-organization of the physiological substrate with the symbolic/linguistic content evolves with respect to changes in vocal effort could be a fruitful extension of this work. For instance, one might ask, is the range of linguistic variability in the sense of variety of grammatical constructions used by the speaker similarly constrained as vocal effort is increased, and is this constraint related to the organization of respiration?

This thesis is obviously only scratching the surface of many interesting questions. At their core, these questions take their cue from the idea that there is a deep and interconnected relationship between physical and symbolic systems. This thesis addressed this issue with respect to structural organization and coordination at the physical level in response to the energetic requirements of projecting linguistic information across a distance. In order to explain biological behavior, the interplay between the physical and the symbolic must be taken into account (Pattee 1978), and the consequences of this interplay must be explored across a multitude of scales. This is an essential problem in the study of complex biological systems generally, and it is *the* essential problem in the study of speech, whether the interplay occurs at the scale of coordination of vocal tract movements in order to produce words, phrases, and sentences, or at the level of the organization and coordination of physiological subsystems to project speech across a distance for the purpose of forming higher-order symbolic relations in social interaction. To put it differently:

Laws of nature do not need embodiments or structures to execute them; rules must have a real physical structure or constraint if they are to be executed...laws hold at all times and all places; rules only exist when and where there are physical structures to execute them. (Pattee 1978, p194).

Coordination involves adapting to the laws of physics in order to form the physical structures that serve as a substrate for the execution of functionally directed action. In the end, this thesis contributes to this more domain-general understanding of coordination by demonstrating the applicability of these principles at a scale of analysis that has received limited attention in speech research. Simply considering the role of respiration in speech motor control allowed this study to find an interesting connection between speech and posture. And more abstractly, it allowed us to directly demonstrate the existence of a systems-level co-dependence between the energetic (i.e. physical) properties of information transmission and physical coordination; or in other words, the ability of the symbolic medium to do actual, physical 'work' by increasing the organization, and thus decreasing the entropy, of the physical substrate supporting its transmission (Deacon 2012). Curiously, this process bears a striking similarity to some recent thinking on communication in the human brain (Nicoleis & Lebedev 2009), where correlation of fluctuations along the physical, neuronal substrate is thought to facilitate the arguably higher order process of communication among groups of non-local neurons, forming the integrated cortical networks that are thought to be crucial to all complex human behaviors. Obviously, at this point, this is nothing more than a striking similarity in need of further exploration. But future work that explicitly addresses the importance of respiration for speech motor coordination will probably have much more to say about this and the other possibilities suggested by the small steps taken in this thesis.

#### REFERENCES

- Abbs, J. H., & Connor, N. P. (1989). Motor coordination for functional human behaviors: Perspectives from a speech motor data base. *Advances in psychology*, *61*, 157-183.
- Ashby, W. R. (1958). Requisite variety and its implications for the control of complex systems. *Cybernetica*, *1*(2), 83-99.
- Bak, P., Tang, C., & Wiesenfeld, K. (1987). Self-organized criticality: An explanation of the 1/f noise. *Physical review letters*, *59*(4), 381.
- Bassingthwaighte, J.B., Liebovitch, S., and West, B.J., (1994) *Fractal Physiology*. Oxford University Press.
- Barbosa, A. V., Yehia, H. C., & Vatikiotis-Bateson, E. (2008). Linguistically valid movement behavior measured non-invasively. In *AVSP* (pp. 173-177).
- Barbosa, A. V., Déchaine, R. M., Vatikiotis-Bateson, E., & Yehia, H. C. (2012). Quantifying timevarying coordination of multimodal speech signals using correlation map analysis. *The Journal of the Acoustical Society of America*, 131, 2162.
- Barbosa, A.V., H. C. Yehia, and E. Vatikiotis-Bateson (2007) MATLAB toolboxfor Audiovisual speech processing, in *Proceedings of the InternationalConference on Auditory-Visual Speech Processing—AVSP 2007*,eds J. Vroomen, M. Swerts, and E. Krahmer (ISCA, Hilvarenbeek, The Netherlands, 2007), pp. 32–37.
- Bernstein, N. A. (1967). The co-ordination and regulation of movements. Pergamon Press Ltd.
- Bertec Corporation. (2009) Bertec Force plate manual.
- Bialas, A., & Koch, V. (1999). Event-by-event fluctuations and inclusive distributions. *Physics Letters B*, 456(1), 1-4.
- Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, 49(3-4), 155-180.
- Buldyrev, S. V., Goldberger, A. L., Havlin, S., Mantegna, R. N., Matsa, M. E., Peng, C. K., & Stanley, H. E. (1995). Long-range correlation properties of coding and noncoding DNA sequences: GenBank analysis. *Physical Review E*, 51(5), 5084.
- Carson, R. G., Goodman, D., Kelso, J. A. S., & Elliott, D. (1995). Phase transitions and critical fluctuations in rhythmic coordination of ipsilateral hand and foot. *Journal of motor behavior*, 27(3), 211-224.
- Chapple, E. D. (1970). Culture and biological man. Holt, Rinehart and Winston.
- Cummins, F. & Port, R.F. (1996). Rhythmic commonalities between hand gestures and speech. In *Proceedings of the eighteenth meeting of the Cognitive Science Society*, pp. 415-419. London: Lawrence Erlbaum

- Daffertshofer, A., Lamoth, C. J., Meijer, O. G., & Beek, P. J. (2004). PCA in studying coordination and variability: a tutorial. *Clinical Biomechanics*, *19*(4), 415-428.
- Dault, M. C., Yardley, L., & Frank, J. S. (2003). Does Articulation Contribute to Modifications of Postural Control During Dual-Task Paradigms? Cognitive Brain Research, 16(3), 434--440.
- Deacon, T. W. (2012). Incomplete nature: How mind emerged from matter. WW Norton & Company.
- Diniz, A., Wijnants, M. L., Torre, K., Barreiros, J., Crato, N., Bosman, A. M., & Delignières, D. (2011). Contemporary theories of 1/ f noise in motor control. *Human movement science*, *30*(5), 889-905.
- Dromey, C., Stathopoulos, E. T., and Sapienza, C. M. (1992). "Glottal airflow and electroglottographic measures of vocal function at multiple intensities," J. Voice 6, 44–54.
- Eke, A., Herman, P., Bassingthwaighte, J., Raymond, G., Percival, D., Cannon, M., & Ikrényi, C. (2000). Physiological time series: distinguishing fractal noises from motions. *Pflügers Archiv*, 439(4), 403-415.
- Eloy, C. (2011). Leonardo's rule, self-similarity, and wind-induced stresses in trees. *Physical review letters*, *107*(25), 258101.
- Frey, U., Silverman, M., Barabasi, A. L., & Suki, B. (1998). Irregularities and power law distributions in the breathing pattern in preterm and term infants. *Journal of Applied Physiology*, 85(3), 789-797.
- Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in cognitive sciences*, *9*(10), 474-480.
- Gelfand, I. M., Gurfinkel, V. S., Tsetlin, M. L., & Shik, M. L. (1971). Some problems in the analysis of movements. In I. M. Gelfand et al. (Eds.), *Models of the structural–functional organization of certain biological systems* (pp. 329–345). Cambridge, MA: MIT Press.
- Giovanni, A., Akl, L., & Ouaknine, M. (2008). Postural dynamics and vocal effort: Preliminary experimental analysis. *Folia Phoniatrica et Logopaedica*, 60(2), 80-85.
- Glass, L. (2001). Synchronization and rhythmic processes in physiology. Nature, 410(6825), 277-284.
- Goodman, L., Riley, M. A., Mitra, S., & Turvey, M. T. (2000). Advantages of rhythmic movements at resonance: minimal active degrees of freedom, minimal noise, and maximal predictability. *Journal of motor behavior*, *32*(1), 3-8.
- Gould W.J. (1971). Effect of respiratory and postural mechanisms upon action of the vocal cords. Folia Phoniatrica. 23: 211–224.
- Gracco, V. L., & Abbs, J. H. (1986). Variant and invariant characteristics of speech movements. *Experimental Brain Research*, 65(1), 156-166.
- Hixon, T. J., Goldman, M.D., & Mead, J. (1973). Kinematics of the chest wall during speech production: Volume displacements of the rib cage, abdomen, and lung. Journal of Speech and Hearing Research, 16, 76-115.

- Holst, F. von (1939/1973). The behavioral physiology of animal and man. Coral Gables, FL: University of Miami Press.
- Horn, B. K., & Schunck, B. G. (1981). Determining optical flow. In *1981 Technical Symposium East* (pp. 319-331). International Society for Optics and Photonics.
- Iwarsson, J., & Sundberg, J. (1999). Breathing behaviors during speech in healthy females and patients with vocal fold nodules. *Logopedics Phonatrics Vocology*, 24(4), 154-169.
- Jeong, B. Y. (1991). Respiration effect on standing balance. Arch Phys Med Rehabil, 72(9), 642-645.
- Jolliffe, I. (1986). Principal Component Analysis. New York: Springer-Verlag
- Kantz, H., & Schreiber, T. (2004). Nonlinear time series analysis (Vol. 7). Cambridge university press.
- Kelso, J. S., Tuller, B., Vatikiotis-Bateson, E., & Fowler, C. A. (1984). Functionally specific articulatory cooperation following jaw perturbations during speech: evidence for coordinative structures. *Journal of Experimental Psychology: Human Perception and Performance*, 10(6), 812.
- Koch, V. (2006). Fluctuations and Correlations in Heavy Ion Collisions. In *Journal of Physics: Conference Series* (Vol. 50, No. 1, p. 95). IOP Publishing.
- Kolasa, J., & Pickett, S. T. A. (1989). Ecological systems and the concept of biological organization. *Proceedings of the National Academy of Sciences*, 86(22), 8837-8841.
- Ladefoged and McKinney. (1963). Loudness, sound pressure, and sub-glottal pressure in speech. *Journal* of the Acoustical Society of America, 35, 454-460
- Lagier, A., Vaugoyeau, M., Ghio, A., Legou, T., Giovanni, A., & Assaiante, C. (2010). Coordination between posture and phonation in vocal effort behavior. *Folia Phoniatrica et Logopaedica*, 62(4), 195-202.
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in psychology*, *4*.
- Lampl, I., Reichova, I., & Ferster, D. (1999). Synchronous membrane potential fluctuations in neurons of the cat visual cortex. *Neuron*, 22(2), 361-374.
- Lashley, K. S. (1951). The problem of serial order in behavior. 1951, 112-135.
- Leonard, T., & Cummins, F. (2011). The temporal relation between beat gestures and speech. *Language* and Cognitive Processes, 26(10), 1457-1471.
- Lenneberg, E. H., Chomsky, N., & Marx, O. (1967). *Biological foundations of language* (Vol. 68). New York: Wiley.
- Liénard, J.-S., and Benedetto, M.-G. D. (1999). "Effect of vocal effort on spectral properties of vowels," J. Acoust. Soc. Am. 106, 411–422.

Little, M., McSharry, P., Moroz, I., & Roberts, S. (2006, May). Nonlinear, biophysically-informed speech pathology detection. In Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on (Vol. 2, pp. II-II). IEEE.

Lombard E (1911) Le signe de l'elevation de la voix. Ann Maladie Oreille Larynx Nez Pharynx 37:101-119.

Mandelbrot, B. B. (1977). Fractals: form, chance and dimension. San Francisco: WH Freeman & Co.

- MATLAB, The MathWorks, Inc., Natick, Massachusetts, United States.
- Munhall, K. G., Jones, J. A., Callan, D. E., Kuratate, T., & Vatikiotis-Bateson, E. (2004). Visual prosody and speech intelligibility head movement improves auditory speech perception. *Psychological science*, *15*(2), 133-137.
- Musha, T., & Yamamoto, M. (1997). 1/f fluctuations in biological systems. In Engineering in Medicine and Biology Society, 1997. Proceedings of the 19th Annual International Conference of the IEEE (Vol. 6, pp. 2692-2697). IEEE.
- Nicolelis, M. A., & Lebedev, M. A. (2009). Principles of neural ensemble physiology underlying the operation of brain-machine interfaces. *Nature Reviews Neuroscience*, *10*(7), 530-540.
- Novikov, E., Novikov, A., Shannahoff-Khalsa, D., Schwartz, B., & Wright, J. (1997). Scale-similar activity in the brain. *Physical Review E*, *56*(3), R2387.
- Pattee, H. H. (1978). The complementarity principle in biological and social structures. *Journal of Social and Biological Structures*, *1*(2), 191-200.
- Peng, C. K., Havlin, S., Stanley, H. E., & Goldberger, A. L. (1995). Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 5(1), 82-87.
- Pettersen, V., Bjørkøy, K., Torp, H., & Westgaard, R. H. (2005). Neck and shoulder muscle activity and thorax movement in singing and speaking tasks with variation in vocal loudness and pitch. *Journal of Voice*, *19*(4), 623-634.
- Pike K.L. (1967) *Language in relation to a unified theory of the structure of human behavior* (Mouton, The Hague).
- Pikovsky, A., Rosenblum, M., & Kurths, J. (Eds.). (2003). Synchronization: a universal concept in nonlinear sciences (Vol. 12). Cambridge university press.
- Pozzo, T., Berthoz, A., & Lefort, L. (1990). Head stabilization during various locomotor tasks in humans. *Experimental Brain Research*, 82(1), 97-106.
- Rajaraman, A., & Ullman, J. D. (2012). Mining of massive datasets. Cambridge University Press.
- Ramig, L. O., Sapir, S., Countryman, S., Pawlas, A. A., O'brien, C., Hoehn, M., & Thompson, L. L. (2001). Intensive voice treatment (LSVT®) for patients with Parkinson's disease: a 2 year follow up. *Journal of Neurology, Neurosurgery & Psychiatry*, 71(4), 493-498.

- Rimmer, K. P., Ford, G. T., & Whitelaw, W. A. (1995). Interaction between postural and respiratory control of human intercostal muscles. *Journal of Applied Physiology*, 79(5), 1556-1561.
- Rusiewicz, H. L. (2010). *The role of prosodic stress and speech perturbation on the temporal synchronization of speech and deictic gestures* (Doctoral dissertation, University of Pittsburgh).
- Sagués, F., Sancho, J. M., & García-Ojalvo, J. (2007). Spatiotemporal order out of noise. *Reviews of Modern Physics*, 79(3), 829.
- Saltzman, E. L., & Munhall, K. G. (1989). A dynamical approach to gestural patterning in speech production. *Ecological psychology*, *1*(4), 333-382.
- Schmidt, R. C., Carello, C., & Turvey, M. T. (1990). Phase transitions and critical fluctuations in the visual coordination of rhythmic movements between people. *Journal of experimental psychology: human perception and performance*, *16*(2), 227.
- Schmidt, R. C., Shaw, B. K., & Turvey, M. T. (1993). Coupling dynamics in interlimb coordination. *Journal of Experimental Psychology: Human Perception and Performance*, 19(2), 397.
- Seshadri, G., & Yegnanarayana, B. (2009). Perceived loudness of speech based on the characteristics of glottal excitation source. *The Journal of the Acoustical Society of America*, *126*(4), 2061-2071.
- Sharp, J. T., Goldberg, N. B., Druz, W. S., & Danon, J. (1975). Relative contributions of rib cage and abdomen to breathing in normal subjects. *Journal of Applied Physiology*, *39*(4), 608-618.
- Shlens, J. (2014). A tutorial on principal component analysis. arXiv preprint arXiv:1404.1100.
- Stone, M., & Vatikiotis-Bateson, E. (1995). Trade-offs in tongue, jaw, and palate contributions to speech production. *Journal of Phonetics*, 23(1), 81-100.
- Suki, B. (2002). Fluctuations and power laws in pulmonary physiology. *American journal of respiratory and critical care medicine*, *166*(2), 133-137.
- Suki, B., Alencar, A.M., Frey, U., Ivanov, P.C., Buldyrev, S. V., Majumdar, A., & Mishima, M. (2003). Fluctuations, noise and scaling in the cardio-pulmonary system. *Fluctuation and Noise Letters*, 3(01), R1-R25.
- Sulter, A. M., and Wit, H. P. (1996). "Glottal volume velocity waveform characteristics in subjects with and without vocal training, related to gender, sound intensity, fundamental frequency, and age," *Journal of the Acoustical. Society of America* 100, 3360–3373
- Sundberg, J., Fahlstedt, E., and Morell, A. (2005). "Effects on the glottal voice source of vocal loudness variation in untrained female and male voices," J. Acoust. Soc. Am. 117, 879–885.
- Torre, K., Delignieres, D., & Lemoine, L. (2007). 1/f β fluctuations in bimanual coordination: An additional challenge for modeling. *Experimental Brain Research*, 183(2), 225-234.
- Traunmüller, H., and Eriksson, A. (2000). "Acoustic effects of variation in vocal effort by men, women, and children," J. Acoust. Soc. Am. 107, 3438–3451.

- Turvey, M. T. (2007). Action and perception at the level of synergies. *Human movement science*, 26(4), 657-697.
- Varela, F., Lachaux, J. P., Rodriguez, E., & Martinerie, J. (2001). The brainweb: phase synchronization and large-scale integration. *Nature reviews neuroscience*, 2(4), 229-239.
- Vatikiotis-Bateson, E., & Ostry, D. J. (1995). An analysis of the dimensionality of jaw motion in speech. *Journal of Phonetics*, 23(1), 101-117.
- Vatikiotis-Bateson, E., Oberg, M., Barbosa, A. V., McAllister, K., Hermiston, N., & Kurth, R. (2009). Postural entrainment by vocal effort in singing and speech. *Proceedings of ESCOM 2009*. Jyväskylä, Finland
- Voss, R. F., & Clarke, J. (1975). 1/f noise in music and speech. Nature, 258, 317-318.
- Walsh, E. (1992). *Muscles, masses and motion: The physiology of normality, hypotonicity, spasticity and rigidity.* London: Mac Keith Press.
- Ward, L. M., & Greenwood, P. E. (2007). 1/f noise. Scholarpedia, 2(12), 1537.
- Warner, R. M. (1979). Periodic rhythms in conversational speech. Language and Speech, 22(4), 381-396.
- West, B. J. (1990). Physiology in fractal dimensions: error tolerance. *Annals of biomedical engineering*, 18(2), 135-149.
- West, B. J. (2013). Fractal physiology and chaos in medicine (Vol. 16). World Scientific.
- West, B. J., & Shlesinger, M. (1990). The noise in natural phenomena. American Scientist, 40-45.
- Winfree, A. T. (1980). The geometry of biological time. *Berlin, Heidelberg and New York: Springer-Verlag*.
- Winkworth, A. L., Davis, P. J., Adams, R. D., & Ellis, E. (1995). Breathing patterns during spontaneous speech. Journal of Speech and Hearing Research, 38, 124-144.
- Winkworth, A. L., & Davis, P. J. (1997). Speech breathing and the Lombard effect. *Journal of Speech, Language, and Hearing Research, 40*(1), 159.
- Winter, D. A. (1995). Human balance and posture control during standing and walking. *Gait & posture*, *3*(4), 193-214.
- Yehia H.C., Rubin P.E., & Vatikiotis-Bateson E. (1998). Quantitative association of vocal-tract and facial behavior. *Speech Communication* 26:23-44.
- Yehia H.C., Kuratate T., & Vatikiotis-Bateson E. (1999) Using speech acoustics to drive facial motion. Proceedings of the 14th International Congress of Phonetic Sciences, eds Ohala JJ, Hasegawa Y, Ohala M, Granville D, & Bailey AC (Linguistics Dept., UC Berkeley), pp 631-634.
- Yehia H.C., Kuratate T., & Vatikiotis-Bateson E. (2002). Linking facial animation, head motion, and speech acoustics. *Journal of Phonetics* 30(3):555-568.

Zahorian, S. A., & Hu, H. (2008). A spectral/temporal method for robust fundamental frequency tracking. *The Journal of the Acoustical Society of America*, *123*(6), 4559-4571.

# **APPENDICES**

## A. Complete statistical results: CMA and SVD

A.1. CMA and SVD within-subjects ANOVA results.

**Table A1.1 Complete within-subjects ANOVA results: CMA and SVD.** Means for all signal pairs in each condition (Normal (N), Loud (L), Louder (Lr), Shout (S), Yell (Y)). ANOVA results for effect of effort condition are listed for m%COR-pos (POS) and m%COR-neg (NEG) measures, as well as for the singular value decomposition analysis (SVD).

			Con	dition M	ANOVA Results				
		N	L	Lr	S	Y	F(4,79)	p <	$\omega_p^2$
	POS	0.051	0.049	0.05	0.053	0.051	0.94	n.s.	-
F0- RMS	NEG	0.042	0.041	0.047	0.036	0.051	8.9	0.0001	0.26
	SVD	206	201	194	178	194	6.1	0.0005	0.19
	POS	0.066	0.067	0.071	0.095	0.075	22.15	0.0001	0.49
F0- Head	NEG	0.066	0.067	0.071	0.096	0.077	23	.0001	0.5
	SVD	114	115	106	92.3	106	21	0.0001	0.47
	POS	0.076	0.076	0.072	0.076	0.066	2.25	n.s.	-
RMS- Head	NEG	0.076	0.076	0.072	0.077	0.068	2.2	n.s.	-
	SVD	114	116	109	100	110	9.1	0.0001	0.27
	POS	0.035	0.037	0.036	0.051	0.058	11.63	0.0001	0.32
F0- Torso	NEG	0.031	0.033	0.035	0.043	0.058	14	0.0001	0.37
	SVD	272	244	232	202	186	7.1	0.0001	0.21
	POS	0.044	0.047	0.046	0.052	0.069	7.00	0.0001	0.21
RMS- Torso	NEG	0.032	0.034	0.034	0.033	0.05	5.3	0.001	0.16
	SVD	220	205	198	180	172	6.6	0.0005	0.2
	POS	0.078	0.083	0.083	0.1	0.11	6.60	0.001	0.20
Torso- Head	NEG	0.078	0.083	0.084	0.1	0.12	6.7	0.0005	0.2
	SVD	113	108	101	91.6	99.2	6.8	0.0001	0.21

Table A.1.1 continued.

			Con	dition M	eans		ANOVA Results		
		N	L	Lr	S	Y	F(4,79)	p <	$\omega_p^2$
	POS	0.18	0.17	0.19	0.2	0.18	4.37	0.01	0.13
COP- Head	NEG	0.18	0.18	0.19	0.2	0.18	3	0.05	0.083
	SVD	85.1	85.3	77.2	72.8	80.8	11	0.0001	0.3
	POS	0.073	0.078	0.076	0.097	0.11	6.67	0.001	0.20
COP- Torso	NEG	0.071	0.077	0.076	0.094	0.11	7	.0001	0.21
	SVD	152	134	124	117	138	4.8	0.005	0.14
	POS	0.067	0.067	0.071	0.084	0.072	8.90	0.0001	0.26
F0- COP	NEG	0.066	0.066	0.071	0.082	0.071	7.6	.0001	0.23
	SVD	145	139	129	119	160	13	0.0001	0.36
	POS	0.077	0.077	0.076	0.072	0.067	2.80	0.05	0.075
RMS- COP	NEG	0.075	0.075	0.075	0.069	0.067	2.7	0.05	0.072
	SVD	143	136	129	124	156	9.5	0.0001	0.28
COP1	POS	0.17	0.17	0.19	0.19	0.15	6.41	0.001	0.20
	NEG	0.16	0.17	0.18	0.18	0.15	5	0.005	0.15
	SVD	97	93.1	85.3	81.6	99.2	7.1	0.0001	0.22

## A.2. CMA and SVD HSD results

**Table A.2.1 m%COR-pos HSD Results.** Pairwise Tukey's HSD tests for all effort conditions. P values are rounded to three significant digits.

Measure	Comparison		p <	Measure	Comp	arison	p <
F0-RMS	Normal	Loud	0.834	Torso-Head	Normal	Loud	0.887
	Normal	Louder	0.957		Normal	Louder	0.88
	Normal	Shout	0.877		Normal	Shout	0.00217
	Normal	Yell	1		Normal	Yell	0.00261
	Loud	Louder	0.997		Loud	Louder	1
	Loud	Shout	0.38		Loud	Shout	0.0221
	Loud	Yell	0.973		Loud	Yell	0.0141
	Louder	Shout	0.555		Louder	Shout	0.0231
	Louder	Yell	0.994		Louder	Yell	0.0145
	Shout	Yell	0.97		Shout	Yell	0.872
F0-Head	Normal	Loud	0.988	COP-Head	Normal	Loud	0.912
	Normal	Louder	0.331		Normal	Louder	0.302
	Normal	Shout	9.93E-09		Normal	Shout	0.0452
	Normal	Yell	0.166		Normal	Yell	0.99
	Loud	Louder	0.632		Loud	Louder	0.0489
	Loud	Shout	9.94E-09		Loud	Shout	0.00612
	Loud	Yell	0.286		Loud	Yell	1
	Louder	Shout	1.34E-08		Louder	Shout	0.727
	Louder	Yell	0.772		Louder	Yell	0.501
	Shout	Yell	0.00391		Shout	Yell	0.155
RMS-Head	Normal	Loud	1	COP-Torso	Normal	Loud	0.806
	Normal	Louder	0.438		Normal	Louder	0.949
	Normal	Shout	1		Normal	Shout	0.0014
	Normal	Yell	0.15		Normal	Yell	0.00416
	Loud	Louder	0.399		Loud	Louder	0.996
	Loud	Shout	1		Loud	Shout	0.0231
	Loud	Yell	0.138		Loud	Yell	0.0283
	Louder	Shout	0.656		Louder	Shout	0.00955
	Louder	Yell	0.677		Louder	Yell	0.0154
	Shout	Yell	0.221		Shout	Yell	0.946

Measure	Comparison		p <	p < Measure		arison	p <
F0-Torso	Normal	Loud	0.901	F0-COP	Normal	Loud	1
	Normal	Louder	0.982		Normal	Louder	0.487
	Normal	Shout	7.73E-05		Normal	Shout	5.56E-06
	Normal	Yell	3.08E-05		Normal	Yell	0.876
	Loud	Louder	0.997		Loud	Louder	0.501
	Loud	Shout	0.00108		Loud	Shout	5.99E-06
	Loud	Yell	0.000221		Loud	Yell	0.882
	Louder	Shout	0.000404		Louder	Shout	0.00086
	Louder	Yell	0.000105		Louder	Yell	1
	Shout	Yell	0.581		Shout	Yell	0.0593
PMS Torso	Normal	Loud	0.817	PMS COP	Normal	Loud	1
KW15-10150	Normal	Louder	0.954	RWD-COI	Normal	Louder	0.985
	Normal	Shout	0.173		Normal	Shout	0.203
	Normal	Yell	3 25E-05		Normal	Yell	0.0598
	Loud	Louder	0.996		Loud	Louder	0.986
	Loud	Shout	0.649		Loud	Shout	0.313
	Loud	Yell	0.00035		Loud	Yell	0.0601
	Louder	Shout	0.458		Louder	Shout	0.562
	Louder	Yell	0.000158		Louder	Yell	0.121
	Shout	Yell	0.0198		Shout	Yell	0.729
COD1 COD2	Normal	Loud	0.000				
COPI-COP2	Normal	Louder	0.999				
	Normal	Shout	0.0129				
	Normal	Vell	0.0221				
	Loud	Louder	0.0252				
	Loud	Shout	0.0252				
	Loud	Yell	0.508				
	Louder	Shout	0.994				
	Louder	Yell	0.00993				
	Shout	Yell	0.00937				

## Table A.2.1 Continued.

Measure	Comparison		p < Measure		Comp	arison	p <
F0-RMS	Normal	Loud	0.899	Torso-Head	Normal	Loud	0.923
	Normal	Louder	0.0942		Normal	Louder	0.873
	Normal	Shout	0.0609		Normal	Shout	0.00298
	Normal	Yell	0.0295		Normal	Yell	0.00178
	Loud	Louder	0.0085		Loud	Louder	1
	Loud	Shout	0.283		Loud	Shout	0.023
	Loud	Yell	0.00641		Loud	Yell	0.00838
	Louder	Shout	7.69E-05		Louder	Shout	0.0312
	Louder	Yell	0.557		Louder	Yell	0.0106
	Shout	Yell	0.000176		Shout	Yell	0.791
F0-Head	Normal	Loud	0.985	COP-Head	Normal	Loud	0.966
	Normal	Louder	0.349		Normal	Louder	0.561
	Normal	Shout	9.92E-09		Normal	Shout	0.114
	Normal	Yell	0.104		Normal	Yell	0.989
	Loud	Louder	0.67		Loud	Louder	0.206
	Loud	Shout	9.93E-09		Loud	Shout	0.0307
	Loud	Yell	0.198		Loud	Yell	1
	Louder	Shout	1.11E-08		Louder	Shout	0.743
	Louder	Yell	0.629		Louder	Yell	0.657
	Shout	Yell	0.00465		Shout	Yell	0.249
RMS-Head	Normal	Loud	0.998	COP-Torso	Normal	Loud	0.723
	Normal	Louder	0.252		Normal	Louder	0.805
	Normal	Shout	1		Normal	Shout	0.00145
	Normal	Yell	0.252		Normal	Yell	0.00122
	Loud	Louder	0.412		Loud	Louder	1
	Loud	Shout	0.997		Loud	Shout	0.033
	Loud	Yell	0.341		Loud	Yell	0.0127
	Louder	Shout	0.4		Louder	Shout	0.0238
	Louder	Yell	0.913		Louder	Yell	0.00981
	Shout	Yell	0.297		Shout	Yell	0.812

 Table A.2.2 m%COR-neg HSD results.
 P values are rounded to three significant digits.

Measure	Comparison		p <	Measure	Comparison		p <
F0-Torso	Normal	Loud	0.806	F0-COP	Normal	Loud	1
	Normal	Louder	0.339		Normal	Louder	0.364
	Normal	Shout	0.000683		Normal	Shout	3.12E-05
	Normal	Yell	3.35E-08		Normal	Yell	0.863
	Loud	Louder	0.937		Loud	Louder	0.4
	Loud	Shout	0.0128		Loud	Shout	3.82E-05
	Loud	Yell	3.87E-07		Loud	Yell	0.879
	Louder	Shout	0.0704		Louder	Shout	0.00633
	Louder	Yell	2.52E-06		Louder	Yell	1
	Shout	Yell	0.00767		Shout	Yell	0.124
RMS-Torso	Normal	Loud	0.919	RMS-COP	Normal	Loud	1
	Normal	Louder	0.935		Normal	Louder	1
	Normal	Shout	0.976		Normal	Shout	0.162
	Normal	Yell	0.000184		Normal	Yell	0.166
	Loud	Louder	1		Loud	Louder	1
	Loud	Shout	1		Loud	Shout	0.187
	Loud	Yell	0.00106		Loud	Yell	0.184
	Louder	Shout	1		Louder	Shout	0.204
	Louder	Yell	0.000957		Louder	Yell	0.196
	Shout	Yell	0.00248		Shout	Yell	0.976
COP1-COP2	Normal	Loud	0.995				
	Normal	Louder	0.0322				
	Normal	Shout	0.0379				
	Normal	Yell	0.766				
	Loud	Louder	0.0853				
	Loud	Shout	0.0834				
	Loud	Yell	0.63				
	Louder	Shout	0.989				
	Louder	Yell	0.0374				
	Shout	Yell	0.0287				

## Table A.2.2 Continued.

Measure	Comparison		p <	Measure	Comp	parison	p <
F0-RMS	Normal	Loud	0.879	Torso-Head	Normal	Loud	0.751
	Normal	Louder	0.109		Normal	Louder	0.0143
	Normal	Shout	0.000138		Normal	Shout	0.000134
	Normal	Yell	0.449		Normal	Yell	0.16
	Loud	Louder	0.551		Loud	Louder	0.254
	Loud	Shout	0.00212		Loud	Shout	0.00416
	Loud	Yell	0.788		Loud	Yell	0.512
	Louder	Shout	0.0777		Louder	Shout	0.285
	Louder	Yell	0.999		Louder	Yell	0.995
	Shout	Yell	0.564		Shout	Yell	0.876
F0-Head	Normal	Loud	0.979	COP-Head	Normal	Loud	1
	Normal	Louder	0.00787		Normal	Louder	0.00141
	Normal	Shout	1.01E-08		Normal	Shout	0.000027
	Normal	Yell	0.125		Normal	Yell	0.913
	Loud	Louder	0.00122		Loud	Louder	0.00106
	Loud	Shout	9.95E-09		Loud	Shout	2.07E-05
	Loud	Yell	0.0575		Loud	Yell	0.895
	Louder	Shout	0.000033		Louder	Shout	0.38
	Louder	Yell	0.995		Louder	Yell	0.584
	Shout	Yell	0.0332		Shout	Yell	0.0898
RMS-Head	Normal	Loud	0.94	COP-Torso	Normal	Loud	0.144
	Normal	Louder	0.171		Normal	Louder	0.00478
	Normal	Shout	4.36E-05		Normal	Shout	0.00346
	Normal	Yell	0.659		Normal	Yell	0.795
	Loud	Louder	0.0276		Loud	Louder	0.711
	Loud	Shout	3.63E-06		Loud	Shout	0.381
	Loud	Yell	0.381		Loud	Yell	0.999
	Louder	Shout	0.0227		Louder	Shout	0.94
	Louder	Yell	1		Louder	Yell	0.86
	Shout	Yell	0.272		Shout	Yell	0.619

 Table A.2.3 SVD HSD Results. P values are rounded to three significant digits.

Measure	Comparison		p <	Measure	Comp	parison	p <
F0-Torso	Normal	Loud	0.265	F0-COP	Normal	Loud	0.43
	Normal	Louder	0.0338		Normal	Louder	0.000816
	Normal	Shout	0.000556		Normal	Shout	2.81E-06
	Normal	Yell	0.00133		Normal	Yell	0.209
	Loud	Louder	0.889		Loud	Louder	0.121
	Loud	Shout	0.0817		Loud	Shout	0.000623
	Loud	Yell	0.0474		Loud	Yell	0.0188
	Louder	Shout	0.358		Louder	Shout	0.177
	Louder	Yell	0.164		Louder	Yell	0.00017
	Shout	Yell	0.902		Shout	Yell	1.95E-06
RMS-Torso	Normal	Loud	0.283	RMS-COP	Normal	Loud	0.398
	Normal	Louder	0.0424		Normal	Louder	0.00225
	Normal	Shout	0.000756		Normal	Shout	0.000543
	Normal	Yell	0.00252		Normal	Yell	0.23
	Loud	Louder	0.906		Loud	Louder	0.244
	Loud	Shout	0.0928		Loud	Shout	0.0484
	Loud	Yell	0.0719		Loud	Yell	0.0198
	Louder	Shout	0.367		Louder	Shout	0.798
	Louder	Yell	0.215		Louder	Yell	0.000388
	Shout	Yell	0.939		Shout	Yell	0.000085
COP1-COP2	Normal	Loud	0.734		ļ		
	Normal	Louder	0.00339				
	Normal	Shout	0.00134				
	Normal	Yell	0.905				
	Loud	Louder	0.104				
	Loud	Shout	0.0297				
	Loud	Yell	0.506				
	Louder	Shout	0.872				
	Louder	Yell	0.0255				
	Shout	Yell	0.00784				
	l						

## Table A.2.3 Continued

# **B.** Complete statistical results: DFA

#### B.1. DFA within-subjects ANOVA results

		Со	ndition N	Means		AN	NOVA Res	sults
	Ν	L	Lr	S	Y	F(4,19)	p <	$\omega_p^2$
СОР	1.14	1.14	1.14	1.23	1.14	2.2	0.105	0.14
COP1	1.23	1.19	1.16	1.21	1.13	2.4	0.0856	0.16
COP2	1.15	1.16	1.16	1.18	1.16	0.23	n.s.	-
FO	0.747	0.765	0.742	0.859	0.849	10.2	0.0005	0.56
Flow	0.986	0.987	0.979	1.02	1.04	3.1	0.05	0.22
Head	1.17	1.17	1.19	1.28	1.28	4.4	0.05	0.32
RMS	0.82	0.825	0.794	0.832	0.852	1.3	n.s.	-

Table B.1.1. DFA within-subjects ANOVA results: acoustic and kinematic signals

Table B.1.2 DFA within-subjects ANOVA results: force plate signals.

		Con	dition N	leans		AN	NOVA Re	sults
	N	L	Lr	S	Y	F(4,19)	p <	$\omega_p^2$
Fx-one	1.12	1.15	1.08	1.19	1.09	6.38	0.005	0.43
Fx-two	1.12	1.15	1.1	1.21	1.11	5.35	0.005	0.38
Fy-one	1.01	1.02	1.02	1.11	0.98	6.09	0.005	0.41
Fy-two	0.991	0.995	0.996	1.08	0.984	4.8	0.01	0.34
Fz-one	1.07	1.07	1.08	1.22	1.06	4.41	0.05	0.32
Fz-two	1.07	1.07	1.08	1.22	1.06	4.15	0.05	0.30
Mx-one	1.2	1.21	1.15	1.25	1.17	3.9	0.05	0.29
Mx-two	1.18	1.16	1.15	1.22	1.16	1.6	n.s.	-
My-one	1.26	1.24	1.19	1.27	1.18	2.3	n.s.	-
My-two	1.22	1.2	1.2	1.24	1.09	3.4	0.05	0.25
Mz-one	1.15	1.14	1.14	1.23	1.17	2	n.s.	-
Mz-two	1.09	1.12	1.11	1.18	1.18	3	0.05	0.22

# B.2. DFA HSD results

Measure	Comp	arison	p <	Measure	Comp	arison	p <
F0	Normal	Loud	0.954	RMS	Normal	Loud	0.999
	Normal	Louder	0.999		Normal	Louder	0.757
	Normal	Shout	0.00196		Normal	Shout	0.98
	Normal	Yell	0.00684		Normal	Yell	0.872
	Loud	Louder	0.88		Loud	Louder	0.627
	Loud	Shout	0.00914		Loud	Shout	0.997
	Loud	Yell	0.0283		Loud	Yell	0.943
	Louder	Shout	0.00117		Louder	Shout	0.431
	Louder	Yell	0.00422		Louder	Yell	0.27
	Shout	Yell	0.998		Shout	Yell	0.992
Head	Normal	Loud	1	Flow	Normal	Loud	1
	Normal	Louder	0.984		Normal	Louder	0.996
	Normal	Shout	0.0533		Normal	Shout	0.528
	Normal	Yell	0.118		Normal	Yell	0.114
	Loud	Louder	0.961		Loud	Louder	0.995
	Loud	Shout	0.0395		Loud	Shout	0.55
	Loud	Yell	0.0908		Loud	Yell	0.121
	Louder	Shout	0.144		Louder	Shout	0.337
	Louder	Yell	0.272		Louder	Yell	0.0603
	Shout	Yell	0.999		Shout	Yell	0.807

 Table B.2.1. DFA HSD Results. P values are rounded to three significant digits.

Measure	Comparison		p <	Measure	Comp	oarison	p <
СОР	Normal	Loud	1	COP-one	Normal	Loud	0.922
	Normal	Louder	1		Normal	Louder	0.53
	Normal	Shout	0.191		Normal	Shout	0.998
	Normal	Yell	0.991		Normal	Yell	0.0818
	Loud	Louder	1		Loud	Louder	0.94
	Loud	Shout	0.242		Loud	Shout	0.984
	Loud	Yell	0.976		Loud	Yell	0.31
	Louder	Shout	0.253		Louder	Shout	0.707
	Louder	Yell	0.972		Louder	Yell	0.711
	Shout	Yell	0.109		Shout	Yell	0.138
COP-two	Normal	Loud	0.996				
	Normal	Louder	0.989				
	Normal	Shout	0.93				
	Normal	Yell	1				
	Loud	Louder	1				
	Loud	Shout	0.992				
	Loud	Yell	0.995				
	Louder	Shout	0.997				
	Louder	Yell	0.988				
	Shout	Yell	0.93				

Table B.2.1 Continued

Measure	Comparison		p <	Measure	Comparison		p <
Fx one	Normal	Loud	0.917	Fx two	Normal	Loud	0.875
	Normal	Louder	0.515		Normal	Louder	0.978
	Normal	Shout	0.112		Normal	Shout	0.0329
	Normal	Yell	0.419		Normal	Yell	0.878
	Loud	Louder	0.149		Loud	Louder	0.564
	Loud	Shout	0.425		Loud	Shout	0.199
	Loud	Yell	0.12		Loud	Yell	0.386
	Louder	Shout	0.004		Louder	Shout	0.00976
	Louder	Yell	0.999		Louder	Yell	0.994
	Shout	Yell	0.00382		Shout	Yell	0.00654
Fy one	Normal	Loud	0.998	Fy two	Normal	Loud	1
	Normal	Louder	1		Normal	Louder	1
	Normal	Shout	0.069		Normal	Shout	0.0731
	Normal	Yell	0.22		Normal	Yell	0.525
	Loud	Louder	1		Loud	Louder	1
	Loud	Shout	0.119		Loud	Shout	0.0899
	Loud	Yell	0.139		Loud	Yell	0.466
	Louder	Shout	0.0863		Louder	Shout	0.0971
	Louder	Yell	0.184		Louder	Yell	0.445
	Shout	Yell	0.00086		Shout	Yell	0.00361
Fz one	Normal	Loud	1	Fz two	Normal	Loud	1
	Normal	Louder	0.998		Normal	Louder	0.999
	Normal	Shout	0.023		Normal	Shout	0.0272
	Normal	Yell	1		Normal	Yell	1
	Loud	Louder	1		Loud	Louder	1
	Loud	Shout	0.0289		Loud	Shout	0.0355
	Loud	Yell	0.998		Loud	Yell	0.999
	Louder	Shout	0.0446		Louder	Shout	0.0479
	Louder	Yell	0.987		Louder	Yell	0.996
	Shout	Yell	0.0232		Shout	Yell	0.0319

# Table B.2.1 Continued
Measure	Comparison		p <	Measure	Comparison		p <
Mx one	Normal	Loud	1	Mx two	Normal	Loud	0.986
	Normal	Louder	0.455		Normal	Louder	0.904
	Normal	Shout	0.525		Normal	Shout	0.785
	Normal	Yell	0.403		Normal	Yell	0.806
	Loud	Louder	0.429		Loud	Louder	0.996
	Loud	Shout	0.553		Loud	Shout	0.49
	Loud	Yell	0.379		Loud	Yell	0.97
	Louder	Shout	0.0281		Louder	Shout	0.3
	Louder	Yell	1		Louder	Yell	0.999
	Shout	Yell	0.0279		Shout	Yell	0.233
My one	Normal	Loud	0.992	My two	Normal	Loud	0.995
	Normal	Louder	0.452		Normal	Louder	0.987
	Normal	Shout	1		Normal	Shout	0.967
	Normal	Yell	0.198		Normal	Yell	0.0652
	Loud	Louder	0.709		Loud	Louder	1
	Loud	Shout	0.969		Loud	Shout	0.842
	Loud	Yell	0.371		Loud	Yell	0.13
	Louder	Shout	0.351		Louder	Shout	0.788
	Louder	Yell	0.963		Louder	Yell	0.156
	Shout	Yell	0.145		Shout	Yell	0.0189
Mz one	Normal	Loud	0.998	Mz two	Normal	Loud	0.848
	Normal	Louder	0.998		Normal	Louder	0.909
	Normal	Shout	0.284		Normal	Shout	0.0332
	Normal	Yell	0.999		Normal	Yell	0.34
	Loud	Louder	1		Loud	Louder	1
	Loud	Shout	0.174		Loud	Shout	0.223
	Loud	Yell	1		Loud	Yell	0.865
	Louder	Shout	0.175		Louder	Shout	0.172
	Louder	Yell	1		Louder	Yell	0.797
	Shout	Yell	0.242		Shout	Yell	0.803

## Table B.2.1 Continued

## C. Complete statistical results: PCA

C.1 PCA within-subjects ANOVA results.

**Table C.1.1 PCA eigenvalue spectrum ANOVA results.** The mean cumulative variance explained by principal component reconstructions of the data consisting of a given number of components (e.g. c1 = one component, c2 = two components) is shown as a function of effort condition (Normal (N), Loud (L), Louder (Lr), Shout (S), Yell (Y). Corresponding ANOVA results for effect of effort condition are listed for each level of PCA reconstruction.

	#PCs		Con	dition M	ANOVA Results				
		Ν	L	Lr	S	Y	F(4,79)	p <	$\omega_p^2$
FPL- One	c1	0.54	0.55	0.57	0.58	0.47	1.9	n.s.	0.039
	c2	0.79	0.8	0.81	0.83	0.74	2.9	0.05	0.079
	с3	0.91	0.92	0.92	0.92	0.88	4.5	0.005	0.14
	c4	0.96	0.97	0.97	0.97	0.95	6.1	0.0005	0.19
	c5	0.99	0.99	.99	0.99	0.99	2.3	n.s.	0.056
FPL- Two	c1	0.55	0.56	0.56	0.53	0.48	2.3	n.s.	0.054
	c2	0.8	0.8	0.8	0.77	0.71	5.3	0.001	0.16
	с3	0.91	0.91	0.92	0.9	0.84	9.8	0.0001	0.28
	c4	0.96	0.97	0.97	0.97	0.93	11	0.0001	0.3
	c5	0.99	0.99	.99	0.99	0.98	5.5	0.001	0.17



## C.2 PCA Cumulative variance plots: Individual participants



## C.3 PCA HSD results

**Table C.3.1. HSD results PCA eigenvalue spectrum: FPL one.** Tukey's HSD tests for differences in effort condition means for PCA reconstructions consisting of a given number of principal components (e.g. c1 = one component, c2 = two components). P values are rounded to three significant digits.

Measure	Comparison		p <	Measure	Comparison		p <
c1	Normal	Loud	0.996	c4	Normal	Loud	0.0158
	Normal	Louder	0.791		Normal	Louder	0.0497
	Normal	Shout	0.838		Normal	Shout	0.334
	Normal	Yell	0.325		Normal	Yell	0.24
	Loud	Louder	0.939		Loud	Louder	0.993
	Loud	Shout	0.946		Loud	Shout	0.947
	Loud	Yell	0.225		Loud	Yell	0.00152
	Louder	Shout	1		Louder	Shout	0.995
	Louder	Yell	0.0879		Louder	Yell	0.00344
	Shout	Yell	0.115		Shout	Yell	0.0181
c2	Normal	Loud	0.991	c5	Normal	Loud	0.319
	Normal	Louder	0.683		Normal	Louder	0.492
	Normal	Shout	0.448		Normal	Shout	0.987
	Normal	Yell	0.198		Normal	Yell	0.627
	Loud	Louder	0.912		Loud	Louder	0.998
	Loud	Shout	0.685		Loud	Shout	0.827
	Loud	Yell	0.113		Loud	Yell	0.0961
	Louder	Shout	0.973		Louder	Shout	0.925
	Louder	Yell	0.0331		Louder	Yell	0.141
	Shout	Yell	0.0201		Shout	Yell	0.483
c3	Normal	Loud	0.441				
	Normal	Louder	0.426				
	Normal	Shout	0.815				
	Normal	Yell	0.0434				
	Loud	Louder	1				
	Loud	Shout	0.998				
	Loud	Yell	0.00225				
	Louder	Shout	0.998				
	Louder	Yell	0.00214				
	Shout	Yell	0.0107				

**Table C.3.2. HSD results PCA eigenvalue spectrum: FPL two.** Tukey's HSD tests for differences in effort condition means for PCA reconstructions consisting of a given number of principal components (e.g. c1 = one component, c2 = two components). P values are rounded to three significant digits.

Measure	Comparison		p <	Measure	Comp	Comparison	
c1	Normal	Loud	0.769	c4	Normal	Loud	0.769
	Normal Louder		0.993		Normal	Louder	0.169
	Normal	Shout	0.886		Normal	Shout	0.998
	Normal	Yell	0.0965		Normal	Yell	0.000023
	Loud	Louder	1		Loud	Louder	0.811
	Loud	Shout	0.793		Loud	Shout	0.965
	Loud	Yell	0.069		Loud	Yell	1.45E-06
	Louder	Shout	0.708		Louder	Shout	0.537
	Louder	Yell	0.0531		Louder	Yell	1.07E-07
	Shout	Yell	0.43		Shout	Yell	0.000048
c2	Normal	Loud	1	c5	Normal	Loud	0.978
	Normal	Louder	1		Normal	Louder	1
	Normal	Shout	0.581		Normal	Shout	0.999
	Normal	Yell	0.000952		Normal	Yell	0.000224
	Loud	Louder	1		Loud	Louder	0.979
	Loud	Shout	0.487		Loud	Shout	1
	Loud	Yell	0.000658		Loud	Yell	0.000761
	Louder	Shout	0.506		Louder	Shout	0.999
	Louder	Yell	0.000708		Louder	Yell	0.000228
	Shout	Yell	0.0484		Shout	Yell	0.00144
c3	Normal	Loud	0.939				
	Normal	Louder	0.917				
	Normal	Shout	0.852				
	Normal	Yell	5.8E-06				
	Loud	Louder	1				
	Loud	Shout	0.481				
	Loud	Yell	9.18E-07				
	Louder	Shout	0.444				
	Louder	Yell	7.7E-07				
	Shout	Yell	0.00037				

 Table C.3.3. HSD results: Number of principal components necessary to represent 96% of force plate data variance. P values are rounded to three significant digits.

Measure	Comparison		p <	Measure	Comparison		p <
FPL-One	Normal	Loud	0.138	FPL-Two	Normal	Loud	0.332
	Normal	Louder	0.064		Normal	Louder	0.0972
	Normal	Shout	0.646		Normal	Shout	0.986
	Normal	Yell	0.802		Normal	Yell	0.0663
	Loud	Louder	0.997		Loud	Louder	0.97
	Loud	Shout	0.982		Loud	Shout	0.842
	Loud	Yell	0.105		Loud	Yell	0.00272
	Louder	Shout	0.922		Louder	Shout	0.539
	Louder	Yell	0.066		Louder	Yell	0.000776
	Shout	Yell	0.309		Shout	Yell	0.047