

NOTE TO USERS

This reproduction is the best copy available.

UMI[®]

THEORETICAL AND EXPERIMENTAL STUDIES OF MULTISENSORY
INTEGRATION AS A COUPLED DYNAMICAL SYSTEM.

by
Collins G. Assisi

A Dissertation Submitted to the Faculty of
The Charles E. Schmidt College of Science in Partial
Fulfillment of the Requirements for the Degree of
Doctor of Philosophy.

Florida Atlantic University
Boca Raton, Florida
August 2005

UMI Number: 3180286

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI[®]

UMI Microform 3180286

Copyright 2005 by ProQuest Information and Learning Company.

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Theoretical and Experimental Studies of Multisensory Integration as a Coupled Dynamical System.

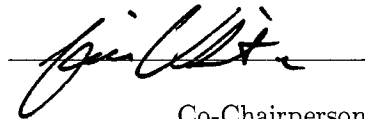
by

Collins G. Assisi

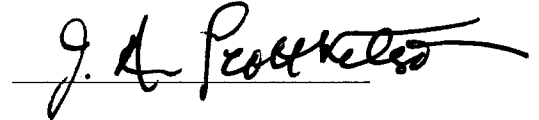
This dissertation was prepared under the direction of the candidate's dissertation co-advisors, Dr. Viktor K. Jirsa and Dr. J. A. Scott Kelso, Program in Complex Systems and Brain Sciences, and has been approved by the members of his supervisory committee.

It was submitted to the faculty of the Charles E. Schmidt College of Science and was accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

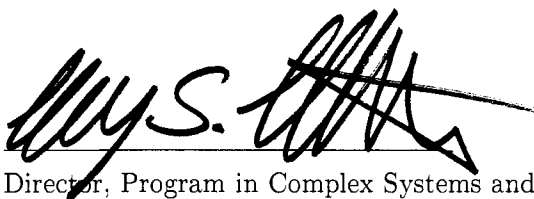
Supervisory Committee



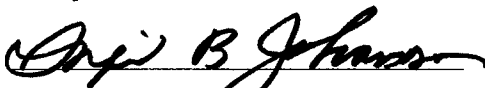
Co-Chairperson



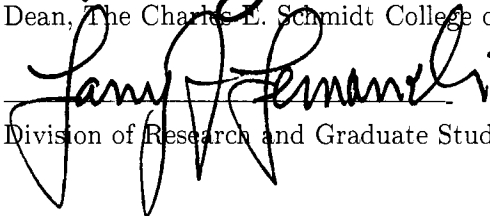
Co-Chairperson



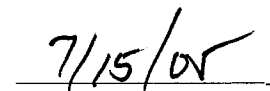
Director, Program in Complex Systems and Brain Sciences



Dean, The Charles E. Schmidt College of Science



Division of Research and Graduate Studies



Date

Abstract

Author: Collins G. Assisi
Title: Theoretical and experimental studies of multisensory integration as a coupled dynamical system.
Institution: Florida Atlantic University
Thesis Advisors: Dr. Viktor K. Jirsa and Dr. J. A. Scott Kelso
Degree: Doctor of Philosophy.
Year: 2005

Perception and behavior are mediated by a widely distributed network of brain areas. Our main concern is, how do the components of the network interact in order to give us a variety of complex coordinated behavior? We first define the nodes of the network, termed functional units, as a strongly coupled ensemble of non-identical neurons and demonstrate that the dynamics of such an ensemble may be approximated by a low dimensional set of equations. The dynamics is studied in two different contexts, sensorimotor coordination and multisensory integration. First, we treat movement coupled to the environment as a driven functional unit. Our central hypothesis is that this coupling must be minimally parametric. We demonstrate the experimental validity of this hypothesis and propose a theoretical model that explains the results of our experiment. A second example of the dynamics of functional units is evident in the domain of multisensory integration. We employ a novel rhythmic multisensory paradigm designed to capture the temporal features of multisensory integration parametrically. The relevant parameters of our experiment are the inter-onset interval between pairs of rhythmically presented stimuli and the frequency of presentation. We partition the two dimensional parameter space using subjects perception of the stimulus sequence. The general features of the partitioning are modality independent suggesting that these features depend on the coupling between the unisensory subsystems. We develop a model with coupled functional units and suggest a candidate coupling scheme. In subsequent chapters we probe the neural correlates of multisensory integration using fMRI and EEG. The results of our fMRI experiment demonstrate that multisensory integration is mediated by a network consisting of primary sensory areas, inferior parietal lobule, prefrontal

areas and the posterior midbrain. Different percepts lead to the recruitment of different areas and their disengagement for other percepts. In analyzing the EEG data, we first develop a mathematical framework that allows us to differentiate between sources activated for both unisensory and multisensory stimulation from those sources activated only for multisensory stimulation. Using this methodology we show that the influences of multisensory processing may be seen at an early (40-60 ms) stage of sensory processing.

Acknowledgements

I am deeply indebted to my advisors Dr. Viktor Jirsa and Dr. Scott Kelso for their guidance and support over the last six years. Viktor has been inspirational in his enthusiasm for science, his willingness to listen and help, his grasp of the larger picture and his tenacious pursuit of the details. I thank him for giving me both, direction, and the space to evolve. I have been fortunate to have Scott as an advisor for the depth of his knowledge, his ability to draw profound relationships between seemingly disparate ideas and his willingness to share his ideas.

I would like to thank Dr. Armin Fuchs and Dr. Steve Bressler for their advice and for serving on my committee.

An environment as scientifically vibrant as the Center would not have been possible without the commitment of the faculty members towards science and teaching. I have greatly benefited from a number of excellent courses taught at the Center and from my interaction with various members of the faculty.

I thank Betty, Armin and KJ for their help in different aspects of this thesis. The motor behavior lab, Gonzalo, Julien, Phil and O2, for all their help with my experiments.

Sam Becker, for his help with collecting the data.

The work in this thesis would not have been possible without the participation of the subjects, who willingly, and sometimes unwillingly, spent long hours staring at blinking led's and hearing interminable beeps, all while twiddling their fingers.

Over the years I have met wonderful people and found great friends at the Center. I would like to thank them for their friendship, their generosity, and their humor that enriched my life. Felix has been occupying the chair next to mine for five years now. I thank him for sharing his thoughts, both profound and profane. Arpan, for his generous friendship, and his delusion that Ganguly will score a century. Ajay, for introducing us to fine single malt, and naming our in-house brand of beer 'immaculate conception'. Murad, for pestering everyone till the abstract notion became 'physical' and we all understood something better. Mukesh, for his child like curiosity and his simplicity. Emmanuelle, for finding the time, the energy, arcane papers under the stone that no one turned, an eye for detail and all the help with my

EEG experiment. It would not get off the ground without your help. Thank you. Momo, for being a great room mate, for letting me steal your groceries, for long heated discussions and for sharing my love for food. Dinesh, for taking the time to organize the food fests.

I thank the students, past and present, for making this place home with your camaraderie and your warmth.

During my first few years at the Center, I spent a great deal of time with three wonderful and interesting people. Shelrie, Cheng and Chris, I thank you for your friendship.

Rhona, for all your help and making sure that the coffee never ran out and the goodies kept coming. We'll win the lottery one of these days.

Denise and Susan for your help with all the paperwork. I'd be broke and lost without you.

If anything fell apart, thank you Bill for putting it back together. Also for sharing your passion for films, chess and Chua circuits.

My friends from Pune for the wonderful times we had together.

Prof. A. D. Gangal, for his patient guidance.

Aparna, for being such a wonderful friend, for expecting so little and giving so much.

Suhita has been part of my life for over a decade now. Thank you for your love and your affection. You've stood by me through everything and I'm lucky to have found you. Nothing would be possible without you.

Mummy and daddy, for their unconditional love and support in everything I've done. I owe it all to you. Linu for being my brother and my friend.

Contents

List of Figures	x
1 Introduction	1
1.1 An approximation of functional units	2
1.2 Dynamics of functional units.	4
1.3 Structure and temporal evolution of the multisensory network.	9
1.3.1 Evidence of multisensory convergence zones	9
1.4 Early interactions in multisensory processing.	11
1.5 Overview	13
2 Synchrony and clustering in heterogeneous networks with global coupling and parameter dispersion.	16
3 Dynamics of multifrequency coordination using parametric driving: Theory and experiment.	26
3.1 Introduction	26
3.2 Background and motivation	28
3.3 Analytical and numerical results	33
3.3.1 Double metronome condition	36
3.3.2 Stochastic properties	38
3.4 Experiment	39
3.4.1 Participants	39
3.4.2 Apparatus	39
3.4.3 Procedure	40

3.4.4	Analysis	40
3.5	Results	42
3.5.1	Frequency	42
3.5.2	Amplitude	46
3.5.2.1	Amplitude-Frequency relationship	46
3.5.3	Phase	49
3.6	Mechanisms of phase transitions	52
3.7	Summary and conclusions	56
4	Asymmetries in temporal multisensory perception.	59
4.1	Introduction	59
4.1.1	Subjects	64
4.1.2	Stimuli	64
4.1.3	Experimental conditions	64
4.1.4	Instructions	65
4.2	Experimental results.	67
4.3	Model	71
4.4	Concluding remarks	79
5	Multisensory integration for timing engages different brain networks: An fMRI study	82
6	Spatial and temporal modulations of multisensory source activity: An EEG study	90
6.1	Introduction	90
6.2	Experimental design	92
6.2.1	Temporal modulations of the unisensory sources due to multisensory stimulation.	95
6.2.2	Time course of uniquely multimodal sources.	96
6.3	Methods	97
6.3.1	Subjects	97

6.3.2	Procedure	97
6.3.3	Data Acquisition	97
6.3.4	Preliminary data analysis	98
6.4	Mathematical formulation of the method of analysis.	98
6.4.1	Relationship between the brain sources and the measured potentials. .	99
6.4.2	Differentiating between the spatial and temporal features of multisensory processing.	100
6.4.3	Adjoint basis decomposition.	102
6.5	Experimental results and discussion.	103
6.5.1	Mode decomposition	104
6.5.2	Comparison between the summative model and the mode decomposition.	105
6.5.3	Early auditory-visual interactions.	109
6.5.4	Uniquely multisensory sources.	109
6.5.4.1	The Drift regime.	109
6.5.4.2	Synchrony vs Asynchrony.	112
6.6	Summary	113
7	Summary	115
A	fMRI materials and methods	118
A.1	Subjects	118
A.2	Behavioral Experiment	118
A.2.1	Experimental design	118
A.2.2	Behavioral Performance	119
A.3	fMRI Experiment	119
A.3.1	Experimental design	119
A.3.2	Image acquisition	120
A.3.3	Data Analysis	120
A.3.4	Tables and legends	121
	Bibliography	123

List of Figures

1.1	A) Shown is the window over which the auditory and visual stimuli cannot be resolved. B) The window of integration depends on the distance between the observer and the event. (Sugita and Suzuki, 2003) C) The visual stimulus is ‘ventriloquised’ towards the auditory stimulus.	6
1.2	Temporal differences between auditory and visual stimuli measured in ms are shown along the axis. Each marked point on the axis is a different estimate of the point where the auditory and the visual stimuli cease to be perceived as simultaneous.	7
2.1	The intrinsic dynamics of a node. The nullclines are in gray. For low values of z_i , (solid cubic nullcline), the trajectory (solid line) settles to a fixed point. Higher values of z_i (dashed cubic nullcline) lead to limit cycle oscillations. (dashed trajectory)	18
2.2	Schematic of different scenarios for a population of coupled neurons in the phase space spanned by the variables \mathbf{x} and \mathbf{y} . Each dot represents the state of a neuron. (a) All the neurons settle to a fixed point. (b) The neurons oscillate in-phase. (c) Some neurons perform sub-threshold oscillations around a fixed point. Others oscillate along a limit cycle. (d) Anti-phase clustering for $K < 0$.	18
2.3	Contour map of the amplitude of the mean field calculated using 100 neurons for different values of the coupling strength K and the parameter dispersion σ . Only positive values of K are shown. See text for details	19
2.4	The timeseries of a population of 100 neurons for various $(K - \sigma)$ values . . .	19

2.5	Contour lines of equal mean field amplitude in $(K - \sigma)$ space. The mean field is calculated from the low-dimensional mode description.	21
2.6	The evolution of the time dependent coefficients, $\xi_1(t)$ (dashed line) and $\xi_2(t)$ (solid line) is plotted for various $(K - \sigma)$ values as a function of time.	22
3.1	Plot of the Arnol'd tongue structures seen for the parametrically driven oscillator system described by equation (3.9). Along the x-axis are the frequencies of the driver Ω and along the y-axis are the strengths of the parametric coupling term ϵ_2 . The lighter regions are Arnol'd tongues. The widest Arnol'd tongue corresponds to a 1:2 coordination mode. The next widest region is the 1:1 mode. Other modes correspond to $\omega/\Omega = k$, where, $k = 3/2, 2, \dots$ are progressively smaller. In the experiment we move from the 1 : 1 to the 1 : 2 coordination mode as the driving frequency Ω is increased along the arrow shown in the figure.	32
3.2	Plot of the amplitude associated with the most dominant frequency of the oscillator ω as the frequency of the driver $\tilde{\Omega}$ is increased. The nonmonotonic behavior of the curve is characteristic of the resonance structure of the oscillator system.	36
3.3	Phase plane trajectories (position vs velocity) obtained from equation (3.9) with an additive Gaussian noise term. a) Single metronome condition: The frequency of the oscillator coincides with the frequency of the metronome. A metronome beat is present only at flexion ($\omega = 2, \tilde{\Omega} = 2$) b) Double metronome condition: The frequency of the oscillator is half that of the metronome. A metronome beat is present at flexion and extension. ($\omega = 2, \tilde{\Omega} = 4$). The other relevant parameters are $\epsilon_1 = 0.1$, $\epsilon_2 = 3$,	38
3.4	Classification of coordination patterns. The smooth curves are the trajectories of finger movement and the rectangular pulses represent the metronome. Amplitude is measured in units of the angular displacement (degrees). (a) An example of a pattern that was classified as 1:1 coordination and (b) Example of a pattern that was classified as 1:2 coordination. See text for details. . . .	41

3.5	The circles show the frequency ratios for five trials at each frequency plateau, and the bars represent the standard deviation(see text for details). ω is the frequency of finger movement and $\tilde{\Omega}$ is the frequency at which the metronome is presented. The data shown in a, b and c are from three subjects and are representative of the patterns seen across all nineteen subjects.	43
3.6	Plotted are the frequency of the driver Ω vs the ratio $\frac{\Omega}{\omega}$ for three possible coupling scenarios between stimulus and movement. The dashed line represents the profile due to purely linear driving obtained by setting $\epsilon_1 = 0.1$ and $\epsilon_2 = 0$. The solid line indicates parametric driving obtained by setting $\epsilon_1 = 0.1$ and $\epsilon_2 = 10$. The dotted line shows the behavior of a the oscillator which, above a maximum threshold frequency, exhibits a constant frequency of oscillation regardless of the frequency of the driver.The other parameter values are from Eq. 3.9 are $A = 1.5; B = 1; \gamma = 0.1; \omega = 4$	44
3.7	Plotted is the width of the 1:1 (empty circles) and 1:2 (filled circles) Arnold tongues as the strength of the parametric coupling term, ϵ_2 , is varied. See text for details.	45
3.8	Each bar corresponds to a subject and the height of the bar corresponds to the width of 1:1 (a) or 1:2 (b) frequency conditions. See text for details. . . .	45
3.9	a) Plot of the mean amplitude as a function of the frequency of the metronome. The mean is obtained for each frequency plateau using data from nineteen subjects. The error bars show the standard deviation across nineteen subjects for all trials. b) The data are from one subject and averaged across five trials. Here, we can see the presence of two peaks at 2.5 Hz and 7 Hz. These features may correspond to the positions of the 1:1and 1:2 Arnol'd tongues	47

3.10	Experimentally obtained phase plane trajectories of finger movement for one subject at frequencies of 4Hz and 8Hz respectively. The boxes indicate the regions of peak flexion and extension. a) At 4Hz (10A) the subject performs a 1:1 mode of coordination and a metronome beat is coincident with peak flexion. b) At 8Hz the subject performs in a 1:2 coordination mode and the metronome is coincident with both peak flexion and extension. c) Standard deviation of the points of peak flexion and peak extension are shown. The white bar shows peak flexion and the black bars show peak extension. Frequency plateaus where the standard deviation of extension is significantly higher than flexion ($p < 0.05$) are marked with an asterisk	48
3.11	a) Standard deviation of the phase of finger movement as a function of the metronome frequency for one subject. Error bars represent the deviation of these values across five trials. b) A plot of the ratio $\frac{\Omega}{\omega}$ vs. Ω for the same subject. See text for details	50
3.12	a) A plot of the phase of finger movement relative to the stimulus onset for one subject. A segment of the data on the border of the 1:1 mode of coordination is chosen. The behavior shows three distinct regions: a 1:1 coordination mode; a region where the drift is initiated; and a region between the 1:1 and 1:2 modes of coordination. Each point corresponds to the phase associated with one cycle of finger movement. b) Plot of the phase of finger movement at points of the movement coincident with a metronome beat obtained from a section of the experimental data for one subject. The data shows phase slippages at $t = 35$ sec and resetting at $t = 40$ sec	51
3.13	Two kind of phase transitions. a) Time series of a coupled oscillator system showing a phase mediated transition. The trajectories are obtained by integrating the HKB model (equations 3.1, 3.2 and 3.3). The transition takes place over a few cycles without a significant change in the amplitude of the oscillation. The solid and the dotted lines are the time series of each oscillator of the coupled oscillator system b) Time series of a parametrically driven oscillator given by equation (3.9) showing an amplitude mediated transition. .	53

3.14	Two mechanisms of phase transitions in rhythmic coordination, an amplitude and a phase mediated transition in the plane spanned by the amplitude, r , and the phase, ϕ	54
3.15	Plot of the time series of the amplitude, r , (a) and the phase, ϕ , (b) during an amplitude mediated transition obtained by simulating Equations 3.19 and 3.20. A drop in the amplitude is accompanied by a jump in the phase through an angle π . $A = 1.5$; $B = 1$; $\gamma = 1$; $\epsilon_1 = 0.1$; $\epsilon_2 = 3$; $\omega = 2$	55
3.16	Plots of the amplitude, r , and the phase, ϕ , as a function of increasing the strength of the linear coupling, ϵ_1 showing the mechanism of an amplitude mediated transition for the equations (3.19,3.20). The solid dots represent the fixed points of the system. The arrows show the approximate direction of the flow in the space spanned by the amplitude, r , and the phase, ϕ . See text for details.	57
4.1	Shown is an example of the stimulus sequence for A (auditory) and V (visual). The parameters of the experiment are dT, the inter-onset interval between two stimuli of a multisensory pair and ISI, the time difference between two consecutively presented stimuli of a particular modality. Similar stimulus sequences were used for Haptic-Visual and Haptic-Auditory stimulation. . . .	63
4.2	a) Shown are the probabilities that subjects perceive AV, VA, S or drift in different parts of the parameter space. The contour lines connect points where the probability that subjects report a given percept are equal. b) probabilities that a given subject perceives HV, VH, S or drift c) probabilities that a given subject perceives HA, AH, S or drift.	68
4.3	The mean width of the S region across subjects for different pairs of stimuli as a function of presentation frequency. The asterisk indicates the widths that are significantly different at a significance level of $p < 0.05$	70

4.4	Shown is a measure of the symmetry of the S region as a function of frequency for the three different stimulus sequences. The black line indicates the number of S responses to the left of the $dT=0$ line and the blue line indicates the number of S responses to the right of the $dT=0$ line a) Auditory-visual stimuli. The black line is significantly higher for frequency = 2Hz and 3.5Hz indicating that the S region tilts towards the AV region. b) The black line indicates the tilt of the S region towards the HV region and the blue line indicates the tilt of S towards VH. c) The black line indicates the tilt of S towards HA and the blue line indicates a tilt towards AH.	72
4.5	Shown is the architecture of the model implemented here. 1 and 2 are the two excitable elements coupled to each other. Temporal delays, τ_1 and τ_2 , are introduced into the coupling between the excitable elements. The delays τ_3 and τ_4 , between the excitable elements and the convergence zone is set to zero. The outputs from the two excitable elements are fed into a convergence zone that classifies the relative dynamics to a prescribed percept. See text for details.	74
4.6	Phase space of the Fitzhugh-Nagumo equations. The dashed lines indicate the x and the y nullclines. The blue lines show different trajectories with different initial conditions. A small perturbation from the fixed point (intersection of the nullclines) causes the system to return after a small deviation. A large perturbation leads it along a long trajectory in phase space. Here the parameters are, $a = 0.7$, $b = 2$, $c = 3$. The parameter b determines the slope of the y nullcline given by the straight line in the figure.	76
4.7	Shown are examples of convergent and divergent dynamics. Each trajectory is associated with an excitable element. The red trajectories indicate that the two elements are coupled while the black trajectories indicate that the two are uncoupled. a) An example of convergence. The trajectories of the two elements come closer to each other for the coupled case than for the uncoupled case. b) The trajectories for the coupled case diverge in comparison to that for the uncoupled case.	77

4.8	Contour plots showing the regions of convergence and divergence in the dT-ISI parameter space. The coupling is instantaneous. $\tau_1 = \tau_2 = 0$ a) Contour map using the parameter $b = 2$, indicating a slow intrinsic dynamics b) Contour map using the parameter $b = 0$, indicating a fast intrinsic dynamics.	80
4.9	The contour map is similar to that shown in figure 4.8. The asymmetry is due to an asymmetric time delay. $\tau_1 = 0$ and $\tau_2 = 7ms$	80
5.1	(A). Experimental design. A visual (V) stimulus is presented after an auditory stimulus (A) a time interval Δt (intra-stimulus interval) and repeated at a rate f (stimulation rate). Behavioral response in the space of Δt and f : the negative values of Δt imply that the visual stimulus (V) precedes the auditory stimulus (A) and the contour levels represent the normalized response (number of responses divided by the total possible responses). Depending on these timing parameters, the participants reported the perceptions of AV (sound before light), VA, S (synchronous) and D (drifting order), thereby partitioning the space into four distinct regions. Notice the asymmetry of S-region, extended more toward the region of AV.	85
5.2	Activations related to crossmodal processing ($p < 0.001$). Areas: inferior frontal (I), superior temporal gyrus (II), middle occipital gyrus (III), inferior parietal lobule (IV), and posterior midbrain (V). The activation in the posterior midbrain (V) in the region of superior colliculus was obtained by a negative contrast of task versus rest. The color intensity represents t-statistics and the activations are overlaid on the Montreal Neurological Institute (MNI) structural template brain in the neurological orientation for display of the t-maps.	86

5.3 (A-C). Activations related with the perception of asynchrony ($p < 0.001$) and the associated mean BOLD responses for the following conditions: (A) ($\Delta t, f$) = (200 ms, 0.5 Hz), (B) (-200 ms, 0.5 Hz) (C) (100 ms, 1.0 Hz). The inferior parietal lobule (IPL) was active as an area of the sub-networks for this percept. The overall signal change between the task and the rest is about 0.3 % ($p < 0.01$). The error bars represent the standard error mean. All subjects were able to establish the perception of asynchrony except for the condition (100 ms, 1.0 Hz), where only nine subjects were able to do that in the scanner. On the other hand, the inferior parietal cortex was not active for (Δt 100 ms, 3.0 Hz), where there was no fixed percept. (D) Activation associated with the perception of synchrony ($p < 0.001$) and the associated mean BOLD response. The negative contrast (18) of the task versus rest activated the posterior midbrain (PM) in the region of superior colliculus for the condition (100ms, 1.0 Hz). The subjects were instructed to get the perception of synchrony and nine subjects could establish the percept. Here, the time series increases continuously during the on-condition upto 6 seconds after the stimuli were turned off and then decreases. The overall signal change between the rest and the task is about 0.3 % ($p < 0.01$). This activation also indicates that the activity in the superior colliculus was correlated with the perception. 87

5.4	Illustrative diagram of the sub-network of active cortical and subcortical brain areas (indicated in red) for crossmodal processing in the perception of asynchrony, synchrony and when there was no clear percept (drift). Based on the brain activations (19) under bimodal conditions, the perception of asynchrony is associated with the sub-network consisting of frontal (F), auditory (A), visual (V), and parietal (P) areas, the perception of synchrony with the sub-network consisting of frontal (F), auditory (A), superior colliculus (Sc) and visual (V) areas. When there was no clear percept formation, active sub-network consisted of frontal (F), auditory (A) and visual (V) areas. (Here, the cortical surface of the MNI brain and MNI single brain template have been used with markers (red and green) to indicate approximate locations for illustrative purposes only).	88
6.1	An example of the stimulus sequence and the parameters of the experiment. Similar stimulus sequences were used for Haptic-Visual and Haptic-Auditory stimulation.	92
6.2	Overview of the behavioral experiment. The contours depict the probability that a subject reports a particular percept for a given point in dT-ISI space. The boxes represent the stimulus conditions that were presented to subjects in the EEG experiment. The red box indicates a condition where we compared the subjects response to two different sets of instructions without changing the physical stimulus (See text for details).	94
6.3	Principal components corresponding to auditory-only and visual-only stimulation for one subject. The number above each figure represents the percentage of variance explained by the corresponding mode. The arrows adjacent to each figure point in the anterior-posterior direction marked as A and P respectively.	105
6.4	Shown is a comparison between $G_{spatial}$ and $G_{summative}$ across subjects for each experimental condition.	107

6.5	Comparison between the summative model (right column) and the time series obtained by projecting $AV(t)$ onto the modes ϕ_A and ϕ_V (left column) and reconstructing the time evolution $AV(t)$	108
6.6	First principal component of the residual time series $R_{AV_{temporal}}$. Each topograph corresponds to one subject. The temporal evolution of the components are shown below the respective topographs. The normalized eigenvalues are shown indicated above each component.	110
6.7	Shown is $P_{spatial}$ as a function of frequency. Each color represents a different subject.	111
6.8	Two different sets of topographs corresponding to the residual signal $R_{SynchAsynch}^2$ is represented by the black trace in panel C of the figure. The topographs in panel A are associated with the first peak of $R_{SynchAsynch}^2$ extending from a latency of 80ms to 150ms. The top two topographs in panel A show different views of the first principal component of $R_{SynchAsynch}$ for the first peak. The bottom two topographs in panel A show different views of the second principal component. The two traces in panel A show the time course of the respective principal components. In panel C the principal components associated with the second peak are shown. The normalized eigenvalues of each component are shown adjacent to the respective component.	114
A.1	Significantly activated brain areas. The anatomical location in MNI coordinates, cluster size ($k \geq 10$ voxels), t and z scores at $p < 0.001$ (uncorrected) for all bimodal conditions. Activations listed with negative contrasts are listed with a negative sign.	122

Chapter 1

Introduction

The issue of functional localization in the brain was a matter of long standing debate, divided between strict cortical localization and holism as diametrically opposite positions (Finger, 2001). A resolution of this controversy came through a redefinition of the notion of *function* (See Luria (1980) for a detailed account). *Function* is usually interpreted to mean one of two things: the *function* of a particular organ such as liver cells, for example, is to secrete bile. This conception of *function* implies a static relationship, a direct immutable causal connection, between the organ or group of cells and what it does. In psychology and neurophysiology *function* has taken a second meaning– to quote Luria “...it is based on a complex dynamic constellation of connections, situated at different levels of the nervous system, that, in the performance of an adaptive task, may be changed with the task itself remaining unchanged”. Large scale cortical networks are a prerequisite for high level functions underlying cognition. Therefore, a major issue that must be addressed is the problem of cortical coordination dynamics (Bressler, 1995; Bressler and Kelso, 2001). How do the components of the network work together to give us a wide variety of complex coordinated behaviors?

The role of large scale brain networks is particularly evident in multisensory integration. Events in the environment are characterized by different physical properties. Sounds traveling at 340 m/s via transverse compressions in the air provide energy to vibrate the tympanic membrane, light falls on the retina, leading to a complex photochemical process that converts light energy to voltage modulations, while physical vibrations impinge on the sensory

receptors of the skin leading to a cascade of processes that gives us our tactile sensation. Each of these sensory receptors are tuned to a particular form of energy and no other. However, our experience of the external world is a coherent, multisensory experience. The brain combines these different energies and identifies it with a particular event or object. This propensity to find a commonality in the simultaneous origins of different physical energies may produce startling illusions (Bertelson and Radeau (1981)). Clearly, the ability of the brain to process information in the form of different sensations into a coherent whole, has significant behavioral advantages. Consider for example, the reduction of the reaction time in response to a multisensory stimulus compared to a unisensory stimulus. Multisensory integration must involve mechanisms that put together information from sensory receptors and brain areas that are widely separated.

In this thesis we will examine different aspects of large scale neuronal networks. The nodes of these networks, termed as a functional units, will be instantiated by a strongly coupled ensemble of neurons. The dynamics of these functional units and their interactions will be examined in the context of sensorimotor coordination and multisensory integration. We wish to address the following questions,

1. What is a mathematically tractable approximation of the functional units?
2. How does the dynamics of functional units manifest itself in behavior and perception ?

1.1 An approximation of functional units

The observables of current brain imaging techniques, like EEG, MEG and local field potentials result from the coordinated activity of ensembles of neurons. These measures do not provide us with detailed information regarding the activity of individual neurons though it highlights their commonality (da Silva, 1991). These measures are known to be indicators of cognitive functions that require the participation of multiple cortical areas. We define a cortical area as a functional unit in the sense of Bressler and Kelso (2001), i.e. as an organized set of locally interacting neuronal populations that receives synaptic inputs and

sends axonal projections. An important question that must be addressed pertains to the dynamics of these functional units. Local populations of neurons are strongly connected within a small volume of cortical tissue and have been approximated as the functional units of cortical processing (Sporns et al., 1989). These volume elements are often described using low dimensional models that may be interpreted as ‘effective neurons’ (Buhmann, 1989). In general a coupled network of dynamical elements may exhibit a rich variety of behavior, ranging from completely synchronized oscillatory and chaotic states (Monte et al., 2003) to clustered oscillations and oscillator death (Strogatz and Mirollo, 1993; Yamaguchi and Shimizu, 1984). The literature on large networks of coupled dynamical systems may be broadly classified into two categories— One deals with the collective behavior of ensembles of **identical** dynamical elements, a simplification that often allows the analysis of complex connection topologies (Pecora and Carrol, 1998). The second approach uses **non-identical** elements at the nodes of the network but is limited to addressing rather simple connection topologies, for example all to all connections or nearest neighbor coupling topologies (Monte et al., 2003; Ott et al., 2002). In general, the neurons that constitute a functional unit possess different intrinsic properties including distributed firing thresholds (Aradi and Soltesz, 2002), different spiking characteristics such as single spikes (Fitzhugh, 1961) and bursts (Hindmarsh and Rose, 1994) among others. Any mathematical model that aims to capture the dynamics of functional units must therefore include non-identical neurons. A tractable mathematical description would require a reduction in the effective dimensionality of the system, that is, though the system may consist of a large number of interacting parts, the collective behavior of the system may be described by a smaller number of equations than would be required to describe all the individual parts. This approach has often been applied to obtain a low dimensional description of globally coupled dynamical systems that are characterized by an all to all connection topology (Kuramoto, 1984). In such networks, also referred to as mean field coupled networks, the summed activity of all the elements is used to drive the activity of each node of the network. Such a low dimensional description would be ideal in describing an ‘effective neuron’. Most traditional approaches have been restricted to studying the synchronized state of the system in addition to small deviations from synchrony (Monte et al., 2003). Synchronous behavior is only one of many possible

collective states of the system. Large deviations from synchronized state such as clustering are seen the collective behavior of coupled dynamical systems and cannot be captured by traditional methods.

A goal of this thesis is to obtain a tractable mathematical description of functional units. This will be the content of Chapter 2 where we propose a method to obtain a low dimensional description of the dynamics of globally coupled networks that include non-identical neurons capable of complex collective behavior. The proposed methodology will be sufficiently general in that it does not assume a specific instantiation of a dynamical system at the nodes (Assisi et al., 2005b). The main contribution of our methodology is its ability to deal with collective behavior that is not synchronous. Given a low dimensional approximation of functional units we then turn to examining the dynamics of these functional units under different conditions. In what follows we will first examine driven functional units in the context of sensorimotor coordination. Second, the phenomenon of multisensory integration will be seen as resulting from the dynamics of interacting functional units.

1.2 Dynamics of functional units.

In order to relate the sensorimotor coordination to the functional units defined above one must first transcend different levels of organization (Jirsa et al., 1998; Kelso et al., 1999; Fuchs et al., 1999). Our experiments concern the rhythmic movement of a finger coordinated with an external auditory metronome, a macroscopic behavioral level of organization (See chapter 3 for details). The functional units are local ensembles of strongly coupled neurons. This level will be termed a neural level of organization. The relationship between the two levels of organization, that is, a mapping between the behavioral and the neural level, is clearly not a simple one. In certain coordination scenarios, especially for the case of rhythmic coordination, brain dynamics (the neural level) may be mapped onto the dynamics of movement (behavioral level) (Fuchs et al., 1992; Kelso et al., 1998). The existence of such a mapping allows us to attribute properties of the functional units to the dynamics of movement. Many aspects of rhythmic movement has been described successfully using limit cycle oscillators, a candidate functional unit. One of the first oscillator models of coordina-

tion behavior is the HKB model (Haken et al., 1985). This model provides a description of experimentally observed bistability and phase transitions in bimanual coordination experiments (Kelso, 1984). In these experiments, subjects were asked to move their index fingers rhythmically in opposite phases (involving simultaneous activation of the flexor and extensor muscles) in time with a metronome. The frequency of the metronome was progressively increased. At a threshold frequency subjects exhibited a spontaneous transition to an in-phase mode of coordination (simultaneous activation of homologous muscle groups). Such transitions have also been noted in a number of other coordination situations and provide a mechanism that affords the system alternate coordination strategies under changing environmental conditions. The effects of the external environment are not explicit in the original HKB model. The coupling between rhythmic movement and an external metronome has been described using linear (Schöner and Kelso, 1988c) and parametric driving terms (Fink et al., 2000; Kay and Warren, 1998; Jirsa et al., 2000). Our experiment, detailed in chapter 3, tests whether the coupling may be described by a parametric driving term. A necessary feature of parametric driving is the existence of coordinative states where finger movement is subharmonically entrained to the metronome. Outside these states known as Arnol'd tongues (see chapter 3 for details) frequency and phase locking between the finger and rhythmic movement is lost. In our analysis of sensorimotor coordination we treat rhythmic finger movement as a driven functional unit. The existence of different states of coordination are not restricted to sensorimotor coordination alone, but has also been noted in experiments on audio-visual perception (Recanzone, 2003; Shipley, 1964).

Thus far we have delved into an example of driven functional units and its relation to sensorimotor coordination. We now consider another example of the dynamics functional units. Specifically, we deal with the interaction between functional units and its relationship to multisensory integration. Multisensory integration is defined as the ability of the brain to combine information from different modalities into a coherent picture of the external world. We will examine the parallels between the dynamics of functional units and multisensory perception. The relationship between the two is particularly evident in the perception of simultaneity.

The perception of simultaneity is an important issue in experimental psychology Spence

and Squire (2003). The problem may be posed as follows– we know that light travels six orders of magnitude faster than sound in air. The initial processing of auditory stimuli is faster than visual processing since audition is not encumbered by sluggish chemical processes like the slow molecular changes in the retinal photopigment. The different temporal delays cancel out when the distance between the multisensory event and the observer is within a radius of ten meters, the so called ‘horizon of simultaneity’. But most multisensory events do not occur at that distance. How then, do we *perceive* events that occur simultaneously as simultaneous? This conundrum lies at the heart of experimental psychology (Levitin et al., 2000; Boring, 1923) because the perception of simultaneity provides us with a particularly salient cue that two different physical energies emanate from the same object or event. That simultaneity is an important cue, is evident from illusions like ventriloquism where the information from two modalities is simultaneous but spatially incongruent (Bertelson and Radeau, 1981). We tend to mislocalize the position of the actual speaker and attribute the voice to a different source. A number of studies has examined the different mechanisms that may provide us with clues into how the brain perceives two events as simultaneous, though they may fall on different sensory organs at different times. Figure 1.1, from Spence and Squire (2003) provides us with an overview of the findings in the behavioral literature on the perception of simultaneity.

The perception of simultaneity in the case of temporally incongruent events has often been attributed to the inability of the brain to resolve two temporally disparate events within a certain window of time. The existence and the nature of this window of simultaneity is a matter of lively debate in the literature. Recently Sugita and Suzuki (2003), based on the results of their behavioral experiments, proposed that the temporal window may be a function of the spatial distance between the event and the observer. This claim has recently found critics in the literature (Arnold et al., 2005) and whether such a distance dependent window exists is still an open question. The width of the window, in particular for auditory-visual stimulation has a long history (Levitin et al., 2000) and has been shown to be notoriously susceptible to variability across subjects and experimental paradigms. Figure 1.2, a representation of the different estimates obtained by various researchers, gives us some clue as to the variability of this window.

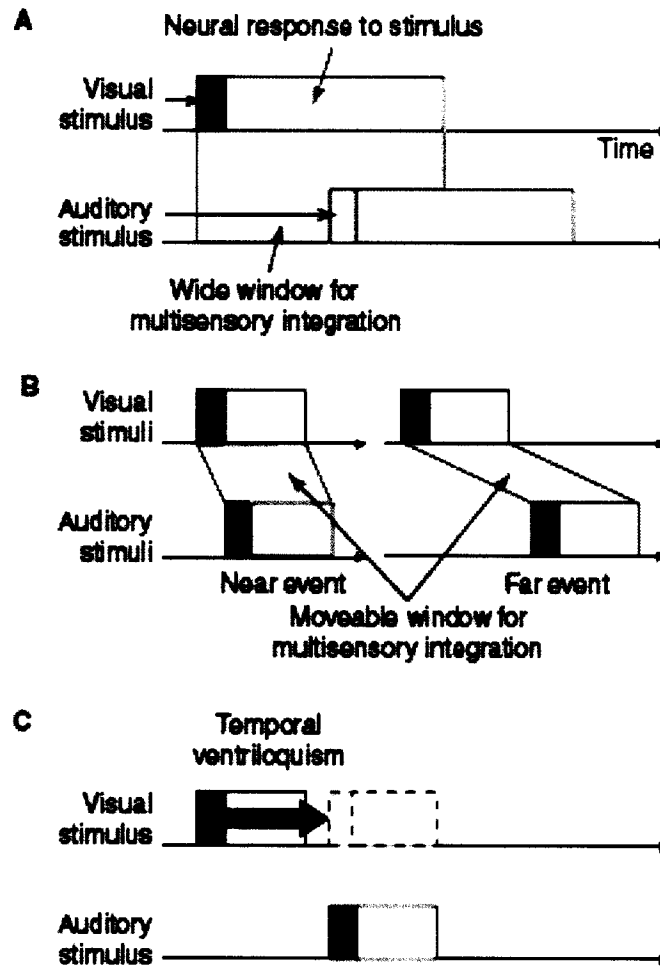


Figure 1.1: A) Shown is the window over which the auditory and visual stimuli cannot be resolved. B) The window of integration depends on the distance between the observer and the event. (Sugita and Suzuki, 2003) C) The visual stimulus is ‘ventriloquised’ towards the auditory stimulus.

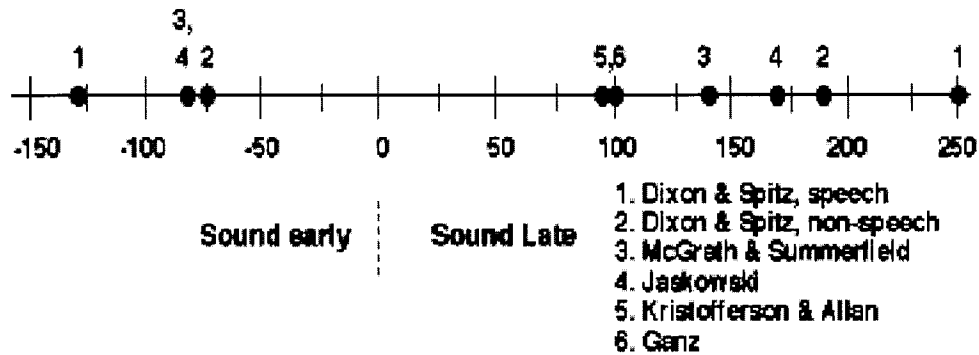


Figure 1.2: Temporal differences between auditory and visual stimuli measured in ms are shown along the axis. Each marked point on the axis is a different estimate of the point where the auditory and the visual stimuli cease to be perceived as simultaneous.

A sharp difference between the estimates relates to the ecological validity of the stimuli presented to subjects. *Ecologically valid* stimuli, for example, audio-visual speech, significantly increase the width of the window. Levitin et al. (2000) argue that human simultaneity perception involves anticipatory mechanisms. Cues suggesting that information in one modality may be followed by another triggers these anticipatory mechanisms. For example, the sight of a falling boulder makes us anticipate the sound of it as it hits the ground. A second, and related mechanism of simultaneity perception, is known in the literature as temporal ventriloquism. Morein-Zemir et al. (2003) suggest that the brain corrects for asynchronous auditory-visual inputs by ‘ventriloquizing’ the perception of the visual stimulus into temporal coincidence with that of the auditory stimulus. Evidence that the presentation of an auditory stimulus affects the perception of visual stimuli has been demonstrated in a number of psychophysical experiments. Shams et al. (2000), in a recent article, demonstrated that presenting two auditory stimuli along with a single visual stimulus led to the perception of two visual stimuli. Shipley (1964) demonstrated that the perceived frequency of a rhythmically presented visual stimulus locks on to an auditory stimulus though the actual frequencies of both the stimuli may be slightly different. An interesting aside to Shipley’s experiment was that beyond a certain frequency the frequency locking phenomenon is lost. Following is his quote describing the subjective experience of the flicker-flutter illusion “... The light and sound were initially pulsating strongly together; as the flutter (auditory)

frequency changed, the apparent flicker (visual) frequency followed right along with it until some point was reached where an apparent asynchrony intruded. It was seldom clear, except at very low frequencies, which sensation, sight or sound, was going faster and which slower, but the rather sudden asynchrony was fairly definite." This emphatic statement about the perception of asynchrony is particularly interesting in the context of our own behavioral experiments and modelling efforts (see chapter 3). The ability of the brain to perceive two related sources of information from different sensory modalities as synchronous and the loss of synchrony beyond certain limits as seen in flicker-flutter illusion has interesting parallels with the dynamics of functional units. In the previous section we noted the existence of multiple states of sensorimotor coordination contingent on the frequency of an external driver. In the flicker-flutter illusion the perception of the auditory stimulus may be seen as driving the visual perception. Beyond a certain frequency difference the frequency locking between the auditory and visual perception of the stimulus is lost akin to the loss of entrainment noted in sensorimotor coordination.

In order to understand the nature of multisensory integration one must first define the coupling between the functional units. There is a great deal of evidence to suggest that multisensory integration is mediated by a large network of brain areas. Some of this literature will be reviewed in the following section. The communication across these widely separated nodes of the network is subject to temporal delays. Typical times for the propagation of an action potential along myelinated axons are approximately 6-9m/s. Time delays in propagation across adjacent gyri separated by distances of approximately 1cm are of the order of 1ms. However, the delays along long axons (20cm) connecting the frontal and occipital regions are of the order of 50-100 ms which is a significant communication delay (Nunez, 1995). In unmyelinated axons the speed of signal conduction is approximately 1m/s (Kandel et al., 1991). A distance of a millimeter between neurons would lead to the conduction time delay of approximately 4% of the period of an oscillation assuming 40Hz oscillations which seems to be an important time scale in feature binding (Tallon-Baudry et al., 1996). Therefore, in trying to understand the mechanisms of multisensory integration, which requires communication across distributed networks, one must account for the temporal scales associated with the individual modalities and the temporal delays

associated with the communication across distinct areas of the brain. We examine this aspect of multisensory integration in detail in Chapter 3. In particular, we will be interested in the asymmetries of temporal scales which may manifest itself as asymmetries in the perception of multisensory stimuli. The existence of multiple intrinsic temporal scales and time delays provides us with additional constraints within which we must examine the interaction of functional units. In what follows we will motivate a plausible architecture to couple these functional units.

1.3 Structure and temporal evolution of the multisensory network.

1.3.1 Evidence of multisensory convergence zones

Current understanding of multisensory integration is motivated by extensive anatomical, neurophysiological and behavioral evidence suggesting a hierarchical model of sensory processing. The hierarchy here implies that simpler features of an object are processed first followed by more complex attributes. An example demonstrating the nature of this hierarchical processing scheme is evident in visual processing (Felleman and Essen, 1991). The receptive fields at successive levels of this hierarchy encode increasingly high level features of the visual stimulus (Hubel and Wiesel, 1965). Receptive fields become larger as the features they represent become progressively more complex. Mesulam (1998) in a recent review suggests that the entire cortical surface may be divided into five distinct zones: i) Primary sensory/motor areas; ii) Unimodal association areas; iii) Multimodal association areas; iv) Limbic areas and v) paralimbic areas. The processing of information starts with transfer from the primary sensory areas to unimodal association areas which receive projections exclusively from modality specific areas of the cortex. Lesions to unimodal association areas are known to result in modality specific deficits. The next stage of processing occurs in heteromodal association areas that receive convergent inputs from more than one modality. Lesions to these areas result in multimodal behavioral deficits. The final limbic zone regulates motivation, memory and emotion. Mesulam's (1998) remarks are of particular interest

in the context of multisensory integration:

“One of the most important principles in the organization of the primate cerebral cortex is the absence of interconnections linking unimodal areas that serve different sensory modalities. . . . many of these unimodal association areas receive monosynaptic feedback projections from heteromodal cortices which are responsive to both auditory and visual stimuli. The sensory-petal (or feedback) projections from heteromodal cortices therefore appear to display a highly selective arrangement that actively protects the fidelity of sensory tuning during the first four synaptic levels of sensory-fugal processing”.

Given the specificity of primary sensory and unimodal association areas (Seltzer and Pandya, 1989), the process of multisensory integration must be attributed to multisensory convergence zones. Neurophysiological and anatomical studies have identified cortical regions in the superior temporal sulcus (Benevento et al., 1977), the parietal cortex (Seltzer and Pandya, 1989), and the frontal cortex (Bodner et al., 1996) that receive inputs from more than one modality. One of the most well studied convergence zones is the superior colliculus (Meredith, 2002; Stein and Meredith, 1993), a laminated midbrain structure. The upper three layers of the superior colliculus receive inputs primarily from the retina and the primary visual cortex. Receptive fields of neurons in this layer are retinotopically organized. The deeper four layers also receive input from extraprimary visual cortex. Neurons in these layers have receptive fields that map the body surface and auditory space (Meredith, 2002; Stein and Meredith, 1993). Each cell in the superior colliculus contains a map of sensory space, one for each modality to which it responds. The receptive fields of maps from different modalities show some degree of spatial overlap. These cells display ‘multisensory integration’ in the following sense: When two or more sensory cues from different modalities appear in close temporal and spatial proximity, the firing rate of these multisensory cells can increase multiplicatively, i.e. beyond that expected by summing the impulses exhibited by each modality. These crossmodal enhancements are maximal when the individual stimuli are minimally effective (a phenomenon called inverse effectiveness, Stein and Meredith (1993)). By contrast, disparate crossmodal cues can produce a profound response depression.

A dominant theme in studying multisensory integration at the cortical level has been to translate these cellular properties to measures that relate to the activity of cortical areas. These cellular properties of multisensory convergence zones translate to the following cri-

teria applied into the local blood oxygen level dependent (BOLD) effect observed in fMRI (Calvert et al., 2001): brain areas show activations when each modality is presented separately. They should show BOLD responses that are significantly greater than the sum of the individual modalities (supra-additivity) during temporally close stimuli and significantly weaker activations during asynchronous stimulation (sub-additivity). Functional imaging studies have also suggested multisensory effects in what are traditionally considered unisensory areas (Calvert et al., 1999). The interactions in these areas have been attributed to feedback processing from multimodal association areas in the hierarchy of sensory processing. In this section we have presented evidence of the existence of regions in the brain, termed as convergence zones, that receive inputs from the unisensory areas and perform the task of integrating these inputs. This provides a plausible architecture to couple multiple functional units, the dynamics of which must then provide us with clues into the nature of multisensory integration.

1.4 Early interactions in multisensory processing.

Thus far we have presented evidence to suggest that multisensory integration is mediated by a widely distributed network of brain regions. In addition, we provide a plausible architecture that subserves multisensory processing consisting of inputs from unisensory areas into convergence zones where information from different modalities is integrated. Recent evidence suggests that the unisensory areas may interact via direct connections, rather than only through feedback connections from the convergence zone. In this section we examine the nature of these early connections. The term *early*, used in describing multisensory interactions is connected to two seemingly related contexts. In the first case *early* is derived from the anatomical literature, where projections between what has traditionally been considered primary sensory and unimodal association areas of distinct sensory modalities have been found. These areas occur *early* within the sensory processing hierarchy. In the second case, the notion of *early* interactions relate to the time at which multisensory processes become evident or the time at which the response to a multisensory stimulus deviates from the response to a unisensory stimulus. There is a great deal of anatomical evidence for projections

between unisensory areas. In addition there is also evidence that the time at which multisensory interactions manifest itself is as early as 40ms. However, it is not yet clear whether the early temporal modifications in the unisensory areas due to multisensory processing can be attributed to the presence of anatomical connections between the unisensory areas.

Over the last six years or so, propelled by some recent anatomical and electrophysiological evidence, the notion of a strict separation between the specific sensory modalities at the early stages of sensory processing has been questioned. In this regard, the work of Falchier et al. (2002) and Rockland and Ojima (2001, 2003) is particularly significant. Falchier et al. (2002) demonstrated the existence of projections from the primary auditory cortex and the polysensory area of the temporal lobe (STP) to the primary visual cortex. These multisensory connections are particularly prominent for that part of V1 subserving the peripheral visual field and are absent for the central visual field. Earlier investigations of the connectivity of area V1 have concentrated primarily on the operculum that subserves the central visual field and may have missed connections to the peripheral visual field. Falchier et al. (2002) posit that these peripheral connections to V1 imply that A1 and STP could participate in a foveation mechanism toward a peripheral sound source. The areas V1 and V2 receive cortical connections not only from visual areas as was previously believed, but also from several parietal and auditory association areas (Rockland and Ojima, 2003, 2001). That these projections have been previously overlooked could be in part due to the inaccessibility of the calcarine fissure. Also, projections, in particular to area V1, are sparse and can be missed because they cannot be distinguished from the background using older tracing techniques (Rockland, 2004). Schroeder et al. (2001) using electrophysiological measures, noticed evidence of somatosensory inputs into the auditory cortex of macaques. They note that the parts of the auditory cortex receiving somatosensory inputs include the caudomedial belt region and may include other belt and parabelt regions but not the primary auditory cortex. These inputs are at the second stage of the auditory processing hierarchy, merely one synapse away from the primary auditory cortex.

That multisensory processes occur early in time has been demonstrated by Giard and Peronnet (1999) and Molholm et al. (2002) in their EEG experiments. Most EEG studies use the summative model of multisensory integration, particularly for the first 200ms post

stimulus. The summative model states that the activity due to multisensory stimulation may be seen as the sum of the activity due to unisensory stimulation in each modality in addition to the activity due to uniquely multisensory processes. Uniquely multisensory processes were attributed to the modulation of the activity of unisensory sources. The reasoning provided by Giard and Peronnet (1999) is that the up to 200ms post stimulus, the scalp potentials are not contaminated by late activities related to target processing or by response selection of motor processes and may thus be regarded as the activity of the unisensory sources. Molholm et al. (2002) using a similar but simpler experimental paradigm used the summative model to suggest that the early auditory-visual interaction effects correspond to both uniquely multisensory sources and the modulation of unisensory activity, in particular early visual sensory processing.

1.5 Overview

In this thesis sensorimotor coordination and multisensory perception are studied as different facets of the dynamics of functional units. In chapter 2 we approximate these functional units by a globally coupled system of non-identical excitable units. Networks with non-identical nodes and global coupling may display a large variety of dynamic behaviors such as phase clustered solutions, synchrony, and oscillator death. The network dynamics is a function of the parameter dispersion and may be captured by conventional mean field approaches if it is close to the completely synchronous state. In this chapter we introduce a novel method based on a mode decomposition in the parameter space, which provides a low-dimensional network description for more complex dynamic behaviors and captures the mean field approach as a special case. The example of globally coupled Fitzhugh-Nagumo neurons is discussed. The low dimensional description thus obtained will be treated as a specific realization of a functional unit. In later chapters we probe the dynamics of these functional units in the context of behavior and perception.

In chapter 3 we study the nature of driven functional units in the context of sensorimotor coordination using a theoretical model and an experimental paradigm designed to test the predictions of the model. Our goal in this chapter is to capture the coupling between the

rhythmic movement of a finger and an external metronome. We demonstrate through a theoretical model and experimental results that the coupling between movement and the external stimulus must be minimally parametric. Different states of coordination of the movement-metronome system are conceived as patterns characteristic of the dynamics of driven functional units.

A second example of the dynamics of functional units is illustrated in chapter 4. Here, we develop a novel experimental paradigm designed to capture the temporal features of multisensory integration parametrically. The relevant parameters of our experiment are the inter-onset interval between pairs of rhythmically presented stimuli and the frequency of presentation. We partitioned this two dimensional parameter space, based on subjects perception of the stimulus sequence into regions of synchrony, asynchrony and a unique multisensory illusion where the perception of the stimulus sequence varies over the length of a trial. We examine the organization and symmetries of this perceptual partitioning for different pairs of stimuli. The general structure of the partitioning is independent of the stimulus modalities involved, but the specific symmetries are modality dependent. This observation suggests that the general features depend on the coupling between the unisensory subsystems. We develop a model consisting of coupled functional units and suggest a candidate coupling scheme that explains the features of the experimental data. In the following two chapters we probe the neural correlates of multisensory integration using fMRI and EEG. In chapter 5 we present the results of an fMRI experiment. The experimental conditions in this experiment are a subset of the behavioral experiment described in Chapter 4. Here, we are primarily interested in obtaining the structure of the network involved in multisensory integration. We demonstrate that this network consists of the primary auditory, the primary visual, inferior parietal lobule, prefrontal areas and the posterior midbrain. Different experimental conditions, that is, different percepts of the stimulus, lead to the selective recruitment of different areas and the disengagement of these areas for other percepts. In chapter 6 we present the results of an EEG experiment. The experimental conditions, as in the earlier fMRI experiment, are a subset of the behavioral experiment described in chapter 4. Here we are primarily concerned with the temporal evolution of networks involved in multisensory integration. The activity measured on the scalp may be attributed to two

factors, i) The activity of the unisensory sources. ii) The neural activity due to uniquely multisensory sources. We first develop a mathematical framework that allows us to differentiate between these two contributions. We then use this newly developed methodology to demonstrate that the influences of multisensory processing may be seen in the putatively unisensory cortices at an early (40-60 ms) in processing. Based on the topography of the residual responses due to uniquely multimodal sources we hypothesize that the generators of these potentials may be in the inferior parietal and frontal areas.

Chapter 2

Synchrony and clustering in heterogeneous networks with global coupling and parameter dispersion.

Low-dimensional dynamics in high-dimensional networks is a ubiquitous phenomenon observed in various physical, chemical and biological problems (Engel et al., 2001; Kuramoto, 1984; Nichols and Wiesenfeld, 1992; Kelso, 1995; Haken, 1983). In particular, networks with dispersed parameters may show such dynamics as a function of the parameter dispersion, including phase clustering and spatiotemporal quasiperiodic and chaotic dynamics (Matthews and Strogatz, 1990). Heterogeneous networks (that is, networks with parameter dispersion) with global coupling have been successfully described using a mean field approach (Monte et al., 2003) which relies on a local expansion around the synchronized solution. Here two sets of reduced equations are derived in which the first describes the dynamics on the synchronization manifold and the second describes the deviations from synchrony. However, because the approach is valid only in the neighborhood of the synchronous solution, global network behaviors such as phase clustering cannot be described low-dimensionally (Rubin and Terman, 2000; Strogatz and Mirollo, 1993). For instance in the field of neuroscience, local populations of neurons are strongly connected within a small volume of cortical tissue and have been interpreted as the functional units of cortical processing (Sporns et al., 1989).

These volume elements are described by a low dimensional ‘effective’ neuron (Buhmann, 1989), identical to the mean field, which will be only legitimate if the local network dynamics is synchronized. Else, conditions are needed which, first, identify when the mean field dynamics fails, and, second, provide alternative approaches to describe the low-dimensional network activity. The aim of our work is to obtain such a general formalism that may be used to approximate the effective dynamics of a high dimensional heterogeneous network of globally coupled elements, which is not limited to synchronous behavior. We demonstrate that the collective behavior of globally coupled dynamical systems with parameter heterogeneity may be captured using a mode decomposition. These modes are defined in the space in which the parameter dispersion occurs and represent distinct patterns of collective behavior. Our approach successfully defines a low-dimensional approximation of the collective dynamics of the system, including various types of asynchronous behavior.

This chapter is organized in two parts. First, we consider a heterogeneous network of globally coupled excitable systems which are often used as toy models for neural behavior. We derive a set of mode equations that approximate its dynamics and, in the process, illustrate the core ideas of our approach. We generalize our formalism to be applicable to networks with arbitrary dynamics at its nodes. One of the simplest examples of an excitable system is the Fitzhugh-Nagumo model (Fitzhugh, 1961). Its variables, x and y operate on slow and fast time scales respectively. A globally coupled network of N Fitzhugh-Nagumo neurons is represented by the following equations

$$\dot{x}_i = c(x_i - \frac{x_i^3}{3} + y_i) + K(X - x_i) + cz_i \quad (2.1)$$

$$\dot{y}_i = \frac{1}{c}(x_i - by_i + a) \quad (2.2)$$

where $i = 1, \dots, N$ and K is the coupling strength. For an uncoupled network, $K = 0$. The intrinsic dynamics of the i -th node is determined by the magnitude of the parameter, z_i , which may be interpreted as an external current or the degree of excitability of the neuronal membrane. For low values of z_i the i -th node has a stable fixed point. If a transient increase in z_i exceeds a threshold, the i -th neuron performs a large excursion in

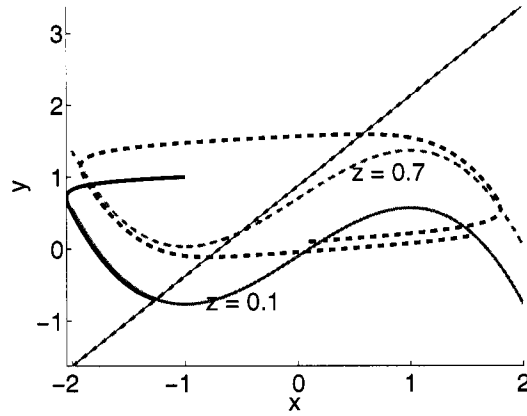


Figure 2.1: The intrinsic dynamics of a node. The nullclines are in gray. For low values of z_i , (solid cubic nullcline), the trajectory (solid line) settles to a fixed point. Higher values of z_i (dashed cubic nullcline) lead to limit cycle oscillations. (dashed trajectory)

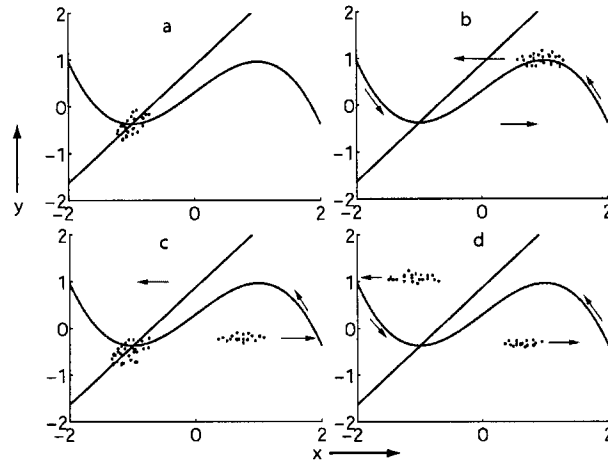


Figure 2.2: Schematic of different scenarios for a population of coupled neurons in the phase space spanned by the variables x and y . Each dot represents the state of a neuron. (a) All the neurons settle to a fixed point. (b) The neurons oscillate in-phase. (c) Some neurons perform sub-threshold oscillations around a fixed point. Others oscillate along a limit cycle. (d) Anti-phase clustering for $K < 0$.

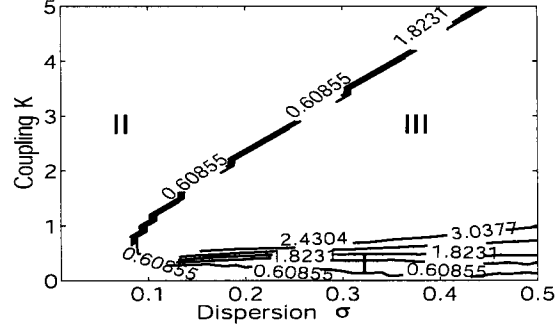


Figure 2.3: Contour map of the amplitude of the mean field calculated using 100 neurons for different values of the coupling strength K and the parameter dispersion σ . Only positive values of K are shown. See text for details

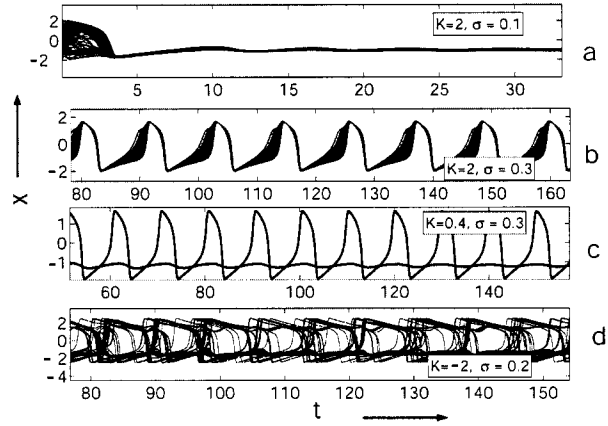


Figure 2.4: The timeseries of a population of 100 neurons for various $(K - \sigma)$ values

phase space before returning to a fixed point representing an action potential. Increasing z_i raises the cubic x -nullcline, destabilizing the fixed point leading to a stable limit cycle (see Fig. 2.1). We introduce heterogeneity in the elements of the network by choosing the parameters z_i from a distribution $g(z)$ with mean μ_z and standard deviation σ . The average activity of the network, $X = \frac{1}{N} \sum_{i=1}^N x_i$, is used to drive each element, implying that every node is connected to every other node and the coupling strength between any two nodes is $\frac{K}{N}$. In dependence of K and the parameter dispersion, σ , we obtain different types of behavior, including a quiescent state (Fig. 2.2a), synchronous in-phase oscillations (Fig. 2.2b), and anti-phase clustering (Fig. 2.2d). In Figure 2.2c, some neurons perform subthreshold oscillations around a fixed point, while others perform supra-threshold oscillations. A characterization of the different types of collective behavior in the parameter space, $K - \sigma$, is shown in Fig. 2.3. Only positive values of K , for which anti-phase clustering is not seen, are shown in the figure. Our numerical simulations show that at each point in the $(K - \sigma)$ space, the mean activity of the population, $X(t)$, is either quiescent or oscillatory. We define the amplitude of the mean field as the difference between the maximum and minimum values of $X(t)$ after the initial transients have settled. Fig. 2.3 shows the contour lines of constant mean field amplitude. Three different regions of mean field dynamics are identified: (I) $K < 0.5$. Here, some of the neurons with low values of z perform subthreshold oscillations around a fixed point, while neurons with higher z values show phase locked supra-threshold oscillations. The dynamics in region I is shown in Fig. 2.2c and Fig. 2.4c. (II) All the neurons settle to a fixed point. (Fig. 2.2a and Fig. 2.4a) (III) The neurons perform supra-threshold synchronous oscillations (Fig. 2.2b and Fig. 2.4b). A transition from region III to region II, mediated by an increase in the coupling strength, K , drives synchronously firing neurons to quiescence. For limit cycle oscillators with dispersed frequencies this transition has been termed oscillator death (Yamaguchi and Shimizu, 1984). An abrupt transition from region II to region III, fixed point behavior to synchronous in-phase oscillations, is mediated by an increase in the parameter dispersion σ . Anti-phase clustered solutions are obtained for $K < 0$. Stable phase clustered solutions have been demonstrated in systems of globally coupled inhibitory neurons (Golomb et al., 1992; Rubin and Terman, 2000), Josephson junctions (Strogatz

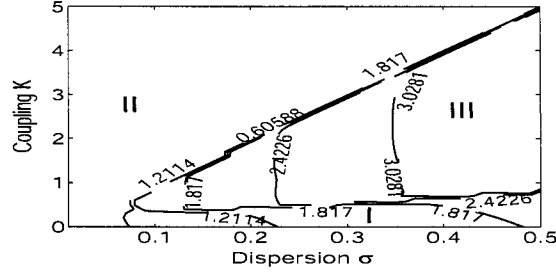


Figure 2.5: Contour lines of equal mean field amplitude in $(K - \sigma)$ space. The mean field is calculated from the low-dimensional mode description.

and Mirollo, 1993) and may be particularly relevant for neurobiological systems (Engel et al., 2001; Kelso, 1995).

The core idea of our approach is the following: The dispersion of the parameter z_i creates an ordering of the nodes with regard to the magnitude of z_i , that is $z_{i+1} \geq z_i \forall i$. We indicate this parametric dependence in the notation of the state vector $\mathbf{x}^i(t)$ of the i -th node by writing $\mathbf{x}^i(t) = (x_i(t), y_i(t))' = (\mathbf{x}_1(z_i, t), \mathbf{x}_2(z_i, t))'$ with the new vector components $\mathbf{x}_k(z_i, t)$, $k = 1, 2$, $i = 1, \dots, N$ parametrized by z_i . The prime denotes the transposed vector. In particular, if the number of neurons is sufficiently large, then it is justified to treat the set $\{z_i\}$, $i = 1, \dots, N$ as a continuous variable $z \in \mathbb{R}$ and the dispersed network state vector $\mathbf{x}(t) = (\dots \mathbf{x}_k(z_i, t) \dots) \longrightarrow \mathbf{x}(z, t)$ as a continuous vector field $\mathbf{x}(z, t) : \mathbb{R}^2 \rightarrow \mathbb{R}$ in the limit for large N . Then the mean field amplitude of the first component which couples the nodes is defined as

$$X(t) = \int_{-\infty}^{\infty} g(z) \mathbf{x}_1(z, t) dz \quad (2.3)$$

where the parameter, z , is distributed according to the function, $g(z)$. The Eq. (2.1) can now be rewritten as

$$\dot{\mathbf{x}}_1(z, t) = c \left(\mathbf{x}_1 - \frac{3}{3} - \mathbf{x}_2 \right) + K [X - \mathbf{x}_1] + cz \quad (2.4)$$

$$\dot{\mathbf{x}}_2(z, t) = \frac{1}{c} (\mathbf{x}_1 - b \mathbf{x}_2 + a) \quad (2.5)$$

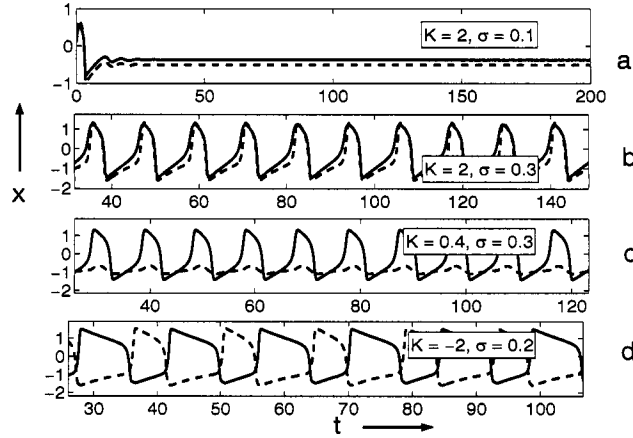


Figure 2.6: The evolution of the time dependent coefficients, $c_1(t)$ (dashed line) and $c_2(t)$ (solid line) is plotted for various $(K - \sigma)$ values as a function of time.

Rewriting the network with heterogeneously distributed parameters as a vector field, $q(z, t)$, allows us to interpret the phenomena due to parameter dispersion as a spatiotemporal pattern formation process in z -space. As a consequence, techniques become available, which allow to decompose the field $q(z, t)$ into its $\leq N$ dominating patterns or modes $\phi_i(z)$ via

$$q(z, t) = \sum_{i=1}^m \begin{pmatrix} \phi_i(t) \\ \psi_i(t) \end{pmatrix} \phi_i(z) + R(z, t) \quad (2.6)$$

with their 2 time-dependent coefficients $\phi_i(t)$, $\psi_i(t)$. The remainder, $R(z, t)$, represents the spatiotemporal dynamics not captured by the modes $\phi_i(z)$. Spatial mode decompositions are not unique and are biased towards certain criteria such as the minimization of the square error (principal component analysis, PCA) or the statistical independence (independent component analysis, ICA). Other techniques utilize a priori knowledge about the system's dynamics to determine interpretable modes (Tesche et al., 1995). If the dispersion has multiple peaks in its distribution $g(z)$, then a separation of the peaks via a mode decomposition in z -space is suggestive. The traditional mean field approach (Monte et al., 2003) is captured by the spatially uniform mode as a special case. In general, the modes will not be orthogonal, but an adjoint basis system $\{\phi_i^\dagger(z)\}$ can always be defined to guarantee

biorthogonality

$$\int_{-\infty}^{\infty} v_i^\dagger(z) v_j(z) dz = \delta_{ij} \quad (2.7)$$

where δ_{ij} is the Kronecker symbol. If the normalized mean square error

$$= \frac{\int_T \int_{-\infty}^{\infty} R(z, t)^2 dz dt}{\int_T \int_{-\infty}^{\infty} q(z, t)^2 dz dt} \quad (2.8)$$

is sufficiently small, then Eq. (2.6) may be truncated after the m -th mode to obtain a low-dimensional description in terms of modes, $m \ll N$.

In our specific examples in Fig. 2.2, we choose the first two principal components, which are approximately non-overlapping rectangular modes, as the spatial modes $v_1(z), v_2(z)$ motivated by the two different collective behaviors (sub- and supra-threshold oscillations). The temporal evolution of the coefficients of the i -th mode, $(\xi_i(t), \eta_i(t))'$, is obtained by projecting Eqs. (2.4, 2.5) onto the mode $v_i(z)$.

$$\dot{\xi}_i(t) = c(\xi_i - \frac{1}{3}\xi_i^3 - \eta_i) + K(\xi_j - \xi_i) + c_i \quad (2.9)$$

$$\dot{\eta}_i = \frac{1}{c}(\xi_i - b\eta_i + a_2) \quad (2.10)$$

where, $i, j = 1, 2$, $\frac{1}{3} = \int_{-\infty}^{\infty} v_1(z)^4 dz$, $\frac{1}{2} = \int_{-\infty}^{\infty} v_1(z) dz$, and $\frac{1}{2} = \int_{-\infty}^{\infty} z v_i(z) dz$. The cross terms resulting from the nonlinearities in Eq. (9) disappear because the modes $v_1(z)$ and $v_2(z)$ do not overlap. The constants c and c_i , modulate the coupling between the modes and are given by

$$c_i = \int_{-\infty}^{\infty} g(z') v_i(z') dz' \cdot \int_{-\infty}^{\infty} v_i(z) dz - 1 \quad (2.11)$$

$$c_i = \int_{-\infty}^{\infty} g(z') v_j(z') dz' \cdot \int_{-\infty}^{\infty} v_i(z) dz \quad (2.12)$$

Eqs. (2.9, 2.10) define the reduced set of mode equations and represent a low-dimensional description of the dispersed network dynamics. These equations are similar to those of two coupled Fitzhugh-Nagumo neurons, justifying the approximation of the population mean activity by “effective” neurons as postulated by (Sporns et al., 1989). To test the reduced

equations computationally, we choose the distribution function, $g(z)$, to be a Gaussian with standard deviation σ and solve the system numerically. Fig. 2.5 describes the resulting different regimes of behavior in analogy to Fig. 2.3: (I) One mode, corresponding to lower values of z , shows subthreshold oscillations while the other mode shows supra-threshold oscillations. (II) Both modes settle to a fixed point. (III) The modes show synchronous in-phase oscillations. A comparison between Figs. 2.3 and 2.5 shows that our low-dimensional mode approximation provides a good description of the population behavior. A goodness of fit, $\text{fit} = 1 - E$, (using Eq. (2.8)) is computed to be 86% for regime (I), 99% for regime (II) and 98% for regime (III). The time series of the modes corresponding to each region in $K - \sigma$ space are shown in Fig. 2.6.

We can now generalize this framework to a globally coupled network with arbitrary dynamical systems at its nodes. Consider a network of N elements given by the following equation

$$\dot{\mathbf{x}}_i = f(\mathbf{x}_i, z_i) + \mathbf{K} \left[\frac{1}{N} \sum_{j=1}^N \mathbf{x}_j - \mathbf{x}_i \right] \quad (2.13)$$

where, $\mathbf{x}_i = [x_i^{(1)} \dots x_i^{(n)}] \in \mathbb{R}^n$ describes the n -dimensional dynamics of the i -th node and $\mathbf{K} \in \mathbb{R}^n \times \mathbb{R}^n$ is a diagonal matrix with nonzero elements of the same value, K . The parameters, z_i , are dispersed according to the distribution function $g(z)$. In the new space defined by z , we introduce the vector field $\mathbf{q}(z, t) = [q_1(z, t) \dots q_n(z, t)]$ where, $q_j(z_i, t) = x_j^{(i)}(t)$, $j = 1, \dots, n$. Then the corresponding dynamics can be written in terms of the new variable $q_j(z, t)$ as follows

$$\dot{q}_j(z, t) = f_j(\mathbf{q}) + K \left[\int_{-\infty}^{\infty} q_j(z, t) g(z) dz - q_j(z, t) \right] \quad (2.14)$$

Assume that the system can be described by a low-dimensional biorthogonal set of m modes in z -space, $\{v_\ell(z), \ell = 1, \dots, m\}$ and its corresponding time-dependent coefficients, $[\xi_\ell^{(1)} \dots \xi_\ell^{(n)}] \in \mathbb{R}^n$. Projecting Eq. (2.14) on each mode $v_\ell^\dagger(z)$, we obtain the following equations for the corresponding time dependent amplitudes

$$\dot{\xi}_\ell^j = f_\ell^j(\{\xi_k\}) + K \left[c_{v_\ell}(\xi_1^j g_{v_1} + \dots + \xi_m^j g_{v_m}) - \xi_\ell^j \right] \quad (2.15)$$

where $h_\ell^j(\{\xi_k\}) = \int_z v_\ell^\dagger \cdot f_\ell^j(\mathbf{q})$ depends on the set $\{\xi_k\}$, $g_{v_\ell} = \int_z v_\ell^\dagger \cdot g(z) dz$ and $c_{v_\ell} = \int_{-\infty}^{\infty} v_\ell^\dagger(z) dz$.

In summary, we have provided a low-dimensional description of a heterogeneous network of globally coupled dynamical systems. We demonstrated the existence of a phase transition from quiescent to synchronous oscillatory behavior mediated by parameter dispersion in a network of Fitzhugh-Nagumo neurons. The effects of parameter dispersion on neuronal systems may be particularly significant. An example can be found in CA1 stratum oriens interneurons in the hippocampus. Evidence from electrophysiological recordings show that hypothermia induced seizures in developing rats causes an increase in the variance of the resting membrane potential without causing a change in the mean resting potential (Aradi and Soltesz, 2002). One of the contributions of this letter is to provide a method to describe phase clustering in networks of dynamical elements. The individual clusters are represented by distinct modes and the interactions among these clusters may be described by the coupling between the mode equations. Further, our method does not assume a particular form of the parameter distribution function and is suitable to study networks with multimodal parameter distributions.

Chapter 3

Dynamics of multifrequency coordination using parametric driving: Theory and experiment.

3.1 Introduction

The coordination of movements in humans and animals may be viewed as ordered spatiotemporal structures that arise in a system consisting of a large number of components and processes organized over different spatial and temporal scales (Kelso et al., 1986; Schöner and Kelso, 1988a). The behavior of the individual components of the system and the spatiotemporal patterns observed at a macroscopic level of description are linked, using the concepts of synergetics, by means of a few observables known as order parameters. These order parameters in turn govern the dynamics at the microscopic level Haken (1983). Over the past two decades or so, using concepts from synergetics and the mathematical tools of nonlinear dynamics, a number of studies has addressed various aspects of the dynamics of coordination, especially rhythmic coordination. Rhythmic movements are archetypes of time dependent behavior in nature. In studies of rhythmic coordination, the elements of the system are often modelled using coupled limit cycle oscillators (Haken et al., 1985; Kelso et al., 1981; Yamanishi et al., 1980; Yuasa and Ito, 1990). External stimuli play an active

role in the dynamics of such limit cycle oscillator systems. Examples range from firefly entrainment (Buck and Buck, 1976) to gait transitions in animals (Collins and Stewart, 1993; Schöner et al., 1990). In the context of biological systems, especially in the area of movement coordination, these can be seen as patterns generated by an action-perception system (Kelso et al., 1990).

One of the first oscillator models of coordination behavior is the HKB model (Haken et al., 1985). This model provides a description of experimentally observed bistability and phase transitions in bimanual coordination experiments (Kelso, 1984). In these experiments, subjects were asked to move their index fingers rhythmically in opposite phases (involving simultaneous activation of the flexor and extensor muscles) in time with a metronome. The frequency of the metronome was progressively increased. At a threshold frequency subjects exhibited a spontaneous transition to an in-phase mode of coordination (simultaneous activation of homologous muscle groups). Such transitions have also been noted in a number of other coordination situations and provide a mechanism that affords the system alternate coordination strategies under changing environmental conditions. Because most of the degrees of freedom of a dynamical system are enslaved close to transition points (Haken, 1983), the system dynamics becomes low-dimensional and hence provides a window into the structure and function of the system. In this chapter we explore different mechanisms that lead to phase transitions in coordination. In recent experiments Fink et al. (2000) investigated the role of environmental information in the dynamics of bimanual coordination. A central finding that emerged from this study was that external information may serve to stabilize states that would otherwise have switched to more stable modes of coordination. Jirsa et al. (2000) accounted for this effect by coupling the external information parametrically to a set of limit cycle oscillators. The main idea elucidated by the above studies is that perception and action, environmental information and the dynamics of movement, are inextricably linked (Kelso et al., 1990). In this paper we present an experimental strategy to further examine the assumption that action and perception are parametrically coupled. Further, based on previous models (Haken et al., 1985; Jirsa et al., 2000; Kay and Warren, 2001; Kelso et al., 1990; Kay and Warren, 1998), we derive a theoretical model for unimanual coordination and compare the features of this model with the observed data.

3.2 Background and motivation

In order to understand the dynamics of the sensorimotor system the following elements must be specified.

1. The dynamics of the uncoupled movement system. After Schöner and Kelso (1988c,b,a), we refer to this as the *intrinsic dynamics*. For our specific experiment, the intrinsic dynamics comprises the rhythmic movement of the finger without reference to the metronome. A number of theoretical and experimental studies have addressed the issue of intrinsic dynamics in different coordination situations. Rhythmic movement of the end effector has been mapped onto the limit cycle of autonomous nonlinear oscillators Kay et al. (1987). This captures the experimental features of reproducible frequency-amplitude-velocity relationships. Kay et al. (1987) have demonstrated the stability of human rhythmic movements through perturbation experiments. Based on these studies a concrete limit cycle model can be formulated.
2. The environmental signal, here a rhythmic metronome. This is operationalized in the present theoretical model using a sinusoidally varying function of time.
3. The coupling between the external stimulus and the movement system. This has been variously described using linear (Schöner and Kelso, 1988c,b) and parametric terms (Jirsa et al., 2000; Kay and Warren, 1998) to drive the limit cycle oscillator. In what follows, we will argue that, given the hypothesis of a limit cycle model and a periodic external signal, a linear driving is insufficient to characterize the dynamics of the system, i.e. the driver must be coupled to the system by higher order terms.

As in Kelso et al. (1990), the present approach is based on a model of bimanual coordination proposed by Haken, Kelso and Bunz in 1985 (Haken et al., 1985). The so-called HKB model consists of two limit cycle oscillators coupled to each other via a nonlinear coupling term. The equations of motion read,

$$\ddot{x}_i + (\tilde{A}x_i^2 + B\dot{x}_i^2 - \tilde{\gamma})\dot{x}_i + \omega_i^2 x_i = f_i(x_1, x_2, \dot{x}_1, \dot{x}_2) \quad (3.1)$$

where,

$$f_1(x_1, x_2, \dot{x}_1, \dot{x}_2) = (\dot{x}_1 - \dot{x}_2)(\gamma + (x_1 - x_2)^2) \quad (3.2)$$

$$f_2(x_1, x_2, \dot{x}_1, \dot{x}_2) = (\dot{x}_2 - \dot{x}_1)(\gamma + (x_1 - x_2)^2) \quad (3.3)$$

x_1 and x_2 represent the movement of the fingers. The system is symmetric under exchange of x_1 and x_2 . The left hand side consists of a linear self-excitation, $\gamma\dot{x}$, a Van der Pol term, $x^2\dot{x}$, and a Rayleigh term, \dot{x}^3 , for saturation. These competing influences define the functional form of the oscillator. The right hand side consists of a nonlinear coupling term depending on x_1 , x_2 and their temporal derivatives. The equations (3.1, 3.2, 3.3) can be reduced to their steady state amplitude and an equation describing the dynamics of the relative phase, $\phi_1 - \phi_2$ (Haken et al., 1985), where ϕ_1 and ϕ_2 are the phases corresponding to x_1 and x_2 , respectively. The amplitude terms are adiabatically eliminated by assuming that the dynamics acts on a time scale much faster than that of the phase variable. The system can then be described by specifying the steady state amplitude, \tilde{A} , and an equation governing the order parameter dynamics, here the dynamics of the relative phase, ϕ ,

$$\dot{\phi} = \frac{\tilde{\gamma}}{\tilde{A} + 3B\omega^2} \quad (3.4)$$

$$\ddot{\phi} = -a \sin \phi - 2b \sin 2\phi \quad (3.5)$$

where, a and b are constants depending on \tilde{A} , \tilde{B} and γ . The phase equation (3.5) has fixed points at 0 and π corresponding to in-phase and anti-phase modes of coordination seen in the experiment. On increasing the frequency of movement, ω , the anti-phase state becomes unstable through a reverse pitchfork bifurcation. Beyond a threshold frequency given by,

$$\omega_c = \sqrt{\frac{4\tilde{\gamma} + \tilde{A}}{-3\tilde{B}}} \quad (3.6)$$

the only stable fixed point corresponds to the in-phase mode of coordination. This is in complete agreement with experimental results. In section 3.6 we will present examples of

situations where an adiabatic elimination of the amplitude is not possible, and consequently, show the existence of two routes via which a transition in the phase can take place. Notice, in equations (3.1,3.2,3.3) the metronome makes only a non-specific contribution to the dynamics of the system, moving the system through various coordinative states via a change in ω , without actively participating in the dynamics. The HKB model does not account for the presence of the metronome and in its original form can be described as a model for the intrinsic dynamics of the system. Recent experiments by Fink et al. (2000) and Byblow et al. (1994) have established that the metronome can modify the trajectories of movement. Further, Fink et al. (2000) demonstrated that the presence of the metronome causes not only local changes in the trajectory of movement but also introduces global effects to the dynamics, such as a shift of the critical frequency at which a phase transition from anti-phase to in-phase modes of coordination takes place. The local changes effected by the metronome, in the context of unimanual movement coordination, will be addressed in subsequent sections.

In light of these experiments it becomes crucial to include the metronome into (3.1). The HKB model correctly reproduces a number of experimentally observed phenomena including phase transitions, hysteresis (Kelso, 1984), critical fluctuations (Kelso et al., 1986) and critical slowing down Scholz et al. (1987) among others. Hence, it is prudent to retain the HKB model at the core of other models that include environmental influences in a description of the dynamics of unimanual (Kelso et al., 1990) and bimanual coordination. For the latter, this was achieved by Schöner and Kelso (1988b) and Jirsa et al. (2000). Schöner and Kelso (1988b) used an additive linear driving term to describe the effect of the metronome. The environmental information was introduced as an additional force acting on the order parameter, ϕ , dynamics attracting it to the phase of the metronome Schöner and Kelso (1988c). In order to account for the results of Fink et al. (2000) recent experiments, Jirsa et al. (2000) used a parametric driving term to describe the effect of the metronome. In the limit of negligible coupling to the metronome both these models reduce to the original HKB equations. Using equations of motion proposed by Jirsa et al. (2000), the coupling function becomes,

$$f_1(x_1, x_2, \dot{x}_1, \dot{x}_2, \epsilon_1, \epsilon_2) = (\dot{x}_1 - \dot{x}_2) \times \\ (\alpha + \beta(x_1 - x_2)^2) + \epsilon_1 \cos \tilde{\Omega}t + \epsilon_2 x_1 \cos \tilde{\Omega}t \quad (3.7)$$

$$f_2(x_1, x_2, \dot{x}_1, \dot{x}_2, \epsilon_1, \epsilon_2) = (\dot{x}_2 - \dot{x}_1) \times \\ (\alpha + \beta(x_1 - x_2)^2) + \epsilon_1 \cos \tilde{\Omega}t + \epsilon_2 x_2 \cos \tilde{\Omega}t \quad (3.8)$$

Where ω is the frequency of the finger movement in (3.1), $\tilde{\Omega}$ is the frequency at which the metronome is presented. ϵ_1 and ϵ_2 are the strengths of the linear and the parametric coupling terms respectively. In our description of unimanual coordination we will use (3.7,3.8). We make the following assumption. The movement of one finger in response to a metronome can be modelled by setting the movement amplitude in one of the oscillators from (3.7,3.8) to zero. The equation of motion of a single finger performing rhythmic movement with a metronome can now be written as follows,

$$\ddot{x}_1 + (Ax_1^2 + B\dot{x}_1^2 - \gamma)\dot{x}_1 + \omega^2 x_1 = \epsilon_1 \cos \tilde{\Omega}t + \epsilon_2 x_1 \cos \tilde{\Omega}t \quad (3.9)$$

where,

$$A = \tilde{A} - \beta, \quad \gamma = \tilde{\gamma} - \alpha \quad (3.10)$$

The crucial element here is the presence of a parametric driving term. Parametric driving leads to characteristic properties of the $\epsilon_2 - \tilde{\Omega}$ parameter space. The parameter space can be divided into stability and instability regions as shown in Figure 3.1.

The frequency of the metronome is plotted along the x -axis and the strength of the parametric coupling term is shown along the y -axis. The lighter regions indicate points in the parameter space where an oscillatory state of a given frequency, that is initially of small amplitude, will grow exponentially. These regions, known as Arnol'd tongues, correspond to coordinative states where the movement of the finger is phase and frequency locked to that of the metronome. Each Arnol'd tongue corresponds to a particular frequency ratio. In the wider tongue, marked 1 : 2, the frequency of the finger, ω , is subharmonically entrained to

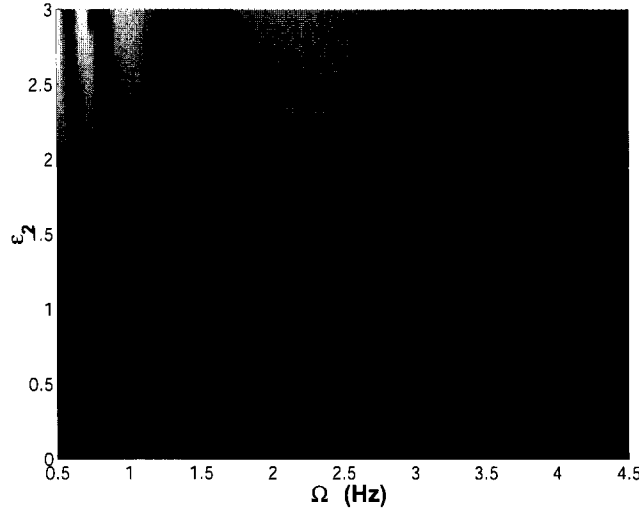


Figure 3.1: Plot of the Arnol'd tongue structures seen for the parametrically driven oscillator system described by equation (3.9). Along the x-axis are the frequencies of the driver Ω and along the y-axis are the strengths of the parametric coupling term ϵ_2 . The lighter regions are Arnol'd tongues. The widest Arnol'd tongue corresponds to a 1:2 coordination mode. The next widest region is the 1:1 mode. Other modes correspond to $\omega : \Omega = k$, where, $k = 3, 2, \dots$ are progressively smaller. In the experiment we move from the 1 : 1 to the 1 : 2 coordination mode as the driving frequency Ω is increased along the arrow shown in the figure.

the frequency of the metronome, $\tilde{\Omega}$, that is, the oscillator completes one cycle for two cycles of the metronome. In the 1 : 1 tongue both the oscillator and the driver have the same frequency. Stable $\omega : \tilde{\Omega}$ (Finger frequency: metronome frequency) coordination modes are obtained for the following frequency ratios (Jirsa et al., 2000).

$$\frac{\omega}{\tilde{\Omega}} = k, \quad k = \frac{1}{2}, 1, \frac{3}{2}, \dots \quad (3.11)$$

Earlier experiments have studied the dynamics of multifrequency coordination (Kelso et al., 1991; Treffner and Turvey, 1996; Peper et al., 1995) and theoretical models of bimanual multifrequency coordination exist (DeGuzman and Kelso, 1991; Haken et al., 1996). Kay and Warren (1998) have described multifrequency coordination using parametric driving to explain the coupling between gait and posture. (see also Kay and Warren (2001))

The goal of our experiment is to drive the system through different coordination regimes, i.e. different frequency ratios by increasing the frequency of the metronome along the direction shown in Figure 3.1. In order to see frequency and phase locking in the experiment

it is necessary that the Arnol'd tongue structures are wide enough so that the system does not fall into a qualitatively different solution due to tiny perturbations that are inevitable in any biological system. In our model, the broadest Arnol'd tongues correspond to the 1:1 and 1:2 modes of coordination and there are no other stable coordination modes in between. Experimentally and operationally, we refer to these as the single and double metronome conditions respectively. In the single metronome condition each cycle of the finger movement corresponds to one cycle of the metronome, whereas in the double metronome condition each cycle of finger movement corresponds to two cycles of the metronome.

3.3 Analytical and numerical results

As our starting point consider (3.9). For the single metronome condition we assume $\omega = \tilde{\Omega}$, i.e. the intrinsic frequency of the finger coincides with that of the metronome. For reasons of mathematical convenience, we write, $\tilde{\Omega} = 2\Omega$, that is, Ω is half the metronome frequency. Here, we are interested in solutions that lie within the 1 : 1 Arnol'd tongue. Consider the following ansatz for the solutions,

$$x(t) = A_0(t) + A_1(t) e^{2i\Omega t} + A_1(t)^* e^{-2i\Omega t} \quad (3.12)$$

The oscillatory system evolves along a closed curve in a two dimensional phase space. The shifts in the center of mass of the closed curve are described by a time dependent real quantity, A_0 . A_1 and A_1^* are time dependent complex quantities that can further be split into a real amplitude, $r \in \mathbb{R}$, and phase, $0 \leq \phi < 2\pi$, as follows,

$$A_1(t) = r(t) e^{i\phi(t)} \quad A_1^*(t) = r(t) e^{-i\phi(t)} \quad (3.13)$$

As a first approximation we can identify two relevant time scales, a fast time scale corresponding to the frequency of the metronome, $\frac{1}{2\Omega}$, and a slow time scale corresponding to the dynamics of the variables, $r(t)$ and $\phi(t)$. Hence, using the slowly varying amplitude approximation, $|\dot{A}| \ll 2\Omega|A|$, we obtain the following equations for the derivatives of (3.12),

$$\dot{x} = 2i\Omega e^{2i\Omega t} + c.c \quad (3.14)$$

and,

$$\ddot{x} = (4i\Omega \dot{A}_1 - 4\Omega^2 A_1) e^{2i\Omega t} + c.c \quad (3.15)$$

Inserting (3.12,3.14,3.15) in (3.9) we obtain terms of the form,

$$e^{2i\Omega t}, e^{-2i\Omega t} \quad (3.16)$$

and,

$$e^{6i\Omega t}, e^{-6i\Omega t} \quad (3.17)$$

Contributions from (17) are negligible due to the rotating wave approximation (Haken, 1983).

By adiabatic elimination, $A_0(t)$ can be expressed in terms of $A_1(t)$ and $A_1^*(t)$.

$$A_0 = \frac{\epsilon_2}{2\omega^2} (A_1 + A_1^*) \quad (3.18)$$

Using these approximations and with the substitutions (3.12,3.13,3.14,3.15) in (3.9) we obtain the following equations for the amplitude and phase of the system,

$$\begin{aligned} \dot{r} = & \left(\frac{\gamma}{2} - \frac{\epsilon_2}{8\omega^2\Omega} \sin 2\phi \right) r - \left\{ \frac{3}{2} \alpha \Omega^2 + \frac{\beta}{2} + \frac{\beta \epsilon_2}{4\omega^4} \right. \\ & \left. (\cos 2\phi + 1) \right\} r^3 - \frac{\epsilon_1}{4\Omega} \sin \phi \end{aligned} \quad (3.19)$$

and

$$\dot{\phi} = \frac{\omega^2 - 4\Omega^2}{4\Omega} - \frac{\epsilon_1}{8\Omega r} \cos \phi - \frac{\epsilon_2^2}{16\Omega\omega^2} (\cos 2\phi + 1) \quad (3.20)$$

The equation (3.20) for the phase consists of contributions from linear and parametric

driving terms. In order to simplify the analysis, we assume that the contribution from the linear driving term is negligible, that is, $\epsilon_1 \ll \epsilon_2$, and $r \neq 0$ because the linear term enforces monostability (see Jirsa et al. (2000)). In this section we consider a situation in which the system is far from a phase transition. A complete description is provided in section 3.6. The steady state value for the phase, ϕ , is governed by the following equation,

$$\cos 2\phi = \frac{4(\omega^2 - 4\Omega^2)\omega^2 - \epsilon_2^2}{\epsilon_2^2} \quad (3.21)$$

which has two steady states corresponding to synchrony of the metronome with peak flexion or peak extension. In order to destabilize one of the solutions, we require a contribution from the linear driving term. In our approximation, we set $\epsilon_1 = 0$, thus neglecting the contribution from the linear driving. In section 3.6 we will elaborate on the different mechanisms of phase transitions when $\epsilon_1 \neq 0$. Note that our experiment is not designed to test this prediction explicitly, which is open to future experimentation. Using (3.21) in (3.19), we obtain a cubic equation of the form,

$$Ar - Br^3 = 0 \quad (3.22)$$

where the coefficients, A and B, are given by,

$$A = \frac{\gamma}{2} - \frac{1}{8\omega\Omega} \sqrt{2(\omega^2 - 4\Omega^2)(\epsilon_2^2 - 2(\omega^2 - 4\Omega^2)\omega^2)} \quad (3.23)$$

and,

$$B = 12\alpha\Omega^2 + \frac{\beta}{2} + \frac{\beta}{\omega^2(\omega^2 - 4\Omega^2)} \quad (3.24)$$

The amplitude is given by,

$$r = \sqrt{\frac{A}{B}} \quad (3.25)$$

Figure 3.2 shows the amplitude response to a change in the frequency of the environmental stimulus ($\tilde{\Omega}$) that drives the oscillator, obtained by integrating Eq. 3.9 for different

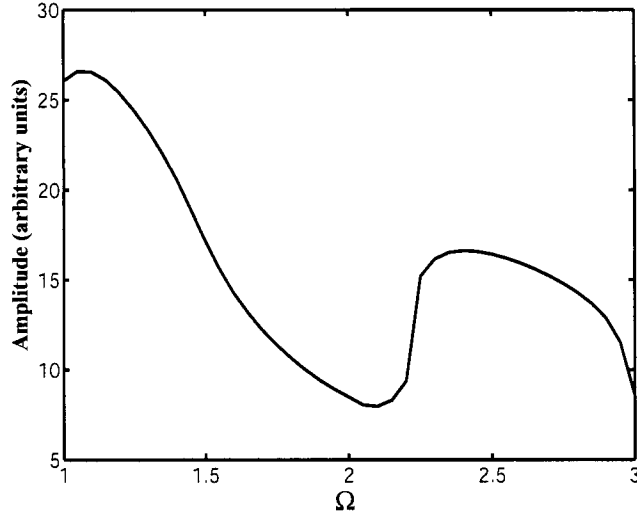


Figure 3.2: Plot of the amplitude associated with the most dominant frequency of the oscillator ω as the frequency of the driver $\tilde{\Omega}$ is increased. The nonmonotonic behavior of the curve is characteristic of the resonance structure of the oscillator system.

values of $\tilde{\Omega}$. The spectral power was computed using a discrete Fourier transform on the steady state time series associated with a particular driving frequency. As the frequency is increased the system moves through the parameter space from the 1 : 1 to the 1 : 2 Arnol'd tongue along the line as shown in Figure 3.1. Within an Arnol'd tongue structure, the amplitude shows an inverted U-shaped dependence with increasing driving frequency. Such nonmonotonic behavior originates due to the finite width of the Arnol'd tongue.

3.3.1 Double metronome condition

Here we seek solutions such that, $\omega = \frac{\tilde{\Omega}}{2} = \Omega$, i.e. the frequency of finger movement is one half that of the metronome. We use the following ansatz for the solutions within the 1 : 2 Arnol'd tongue,

$$x(t) = A_1 e^{i\Omega t} + A_1^* e^{-i\Omega t} \quad (3.26)$$

The analysis follows the lines of the single metronome condition. As before, A_1 and A_1^* are time dependent coefficients that can be split into amplitude and phase components.

$$A_1(t) = r(t)e^{i\phi(t)} \quad A_1^*(t) = r(t)e^{-i\phi(t)} \quad (3.27)$$

Performing the slowly varying amplitude and rotating wave approximations, with (3.26,3.27) in (3.9) we obtain the following equations for the amplitude and phase of the oscillator,

$$\dot{r} = \left(-\frac{3}{2}\alpha\Omega^2 + \frac{\beta}{2}\right)r^3 + \left(\frac{\gamma}{2} - \frac{\epsilon_2}{4\Omega}\sin 2\phi\right)r \quad (3.28)$$

and,

$$\dot{\phi} = \frac{\omega^2 - \Omega^2}{2\Omega} - \frac{\epsilon_2}{4\Omega}\cos 2\phi \quad (3.29)$$

Considering the steady state solutions of (3.28) and (3.29), we have, for ϕ

$$\cos 2\phi = \frac{2(\omega^2 - \Omega^2)}{\epsilon_2} \quad (3.30)$$

Replacing this in the (3.28) we obtain an equation describing the steady state amplitude.

$$Ar - Br^3 = 0 \quad (3.31)$$

where,

$$A = \frac{\gamma}{2} - \frac{1}{4\Omega}\sqrt{\epsilon_2^2 - 4(\omega^2 - \Omega^2)^2} \quad (3.32)$$

and

$$B = \frac{3}{2}\alpha\Omega^2 + \frac{\beta}{2} \quad (3.33)$$

The steady state amplitude is given by,

$$r = \sqrt{\frac{A}{B}} \quad (3.34)$$

As in the single metronome condition, here too, we observe an inverted U-shaped amplitude dependence. In Figure 3.2 the two humps correspond to the single and double

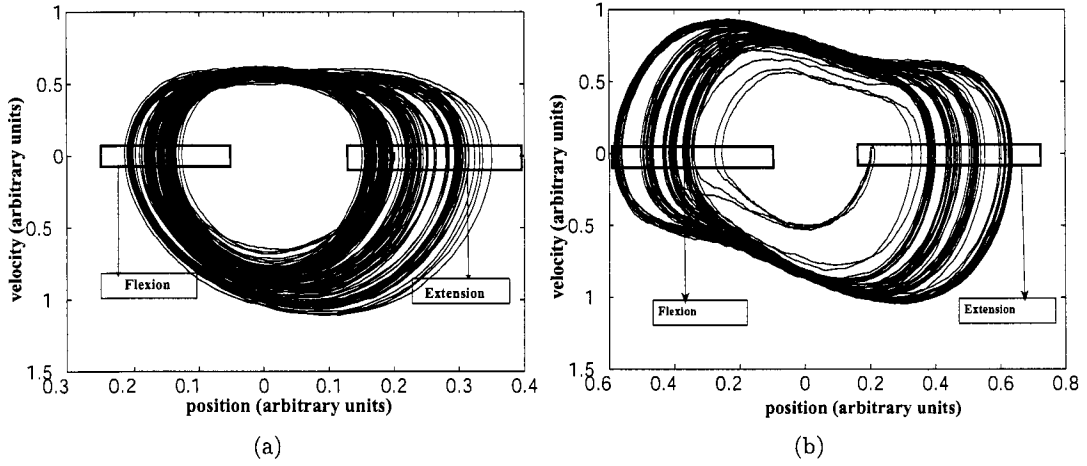


Figure 3.3: Phase plane trajectories (position vs velocity) obtained from equation (3.9) with an additive Gaussian noise term. a) Single metronome condition: The frequency of the oscillator coincides with the frequency of the metronome. A metronome beat is present only at flexion ($\omega = 2, \tilde{\Omega} = 2$) b) Double metronome condition: The frequency of the oscillator is half that of the metronome. A metronome beat is present at flexion and extension. ($\omega = 2, \tilde{\Omega} = 4$). The other relevant parameters are $\epsilon_1 = 0.1$, $\epsilon_2 = 3$,

metronome condition respectively.

3.3.2 Stochastic properties

A characteristic property of all biological systems is the presence of noise, the sources of which are both internal and environmental. The movements are therefore not exactly reproducible across trials or even across cycles of a single trial. In order to account for the variability seen in the experimental data we must include noise into (3.9). The simplest way to do this is by using an additive Gaussian noise, $\xi(t)$, term with the following properties.

$$\langle \xi(t) \rangle = 0, \langle \xi(t) \xi(t + \tau) \rangle = Q^2 \delta(\tau), \quad Q = 0.01 \quad (3.35)$$

Schöner et al. (1986) studied the effects of noise for the case of self-paced bimanual coordination. For a parametrically driven oscillator the trajectories of motion are asymmetric in phase space. The variability is distinctly lower at points of the trajectory that are coincident with peaks of the external signal as compared to other points. In Figures 3.3a and 3.3b we summarize the results of a numerical simulation obtained from adding a Gaussian

noise term to (3.9). Figure 3.3a describes the single metronome condition for a particular driving frequency. The driving frequency is at the center of the 1 : 1 Arnol'd tongue (1.2Hz). The points of peak flexion and extension are marked in the figure. Clearly, the trajectory shows lower variability at peak flexion in comparison to peak extension. This effect can be attributed to the presence of a stimulus at peak flexion which is absent at the opposite reversal point. Figure 3.3b shows the effect of the stimulus on the trajectories of motion for the double metronome condition. A frequency was chosen at the center of the 1 : 2 Arnol'd tongue. Here, unlike in the previous case, Figure 3.3a, the trajectories are symmetric in phase space. In the double metronome condition, the environmental stimulus occurs at both peak flexion and peak extension. Hence, the trajectories at the two reversal points show similar variability. A discussion of these results will follow in the next section where we compare the theoretical predictions of the model with our experimental observations.

3.4 Experiment

In the experiment subjects moved their index finger back and forth on a parasagittal plane in accord with the frequency of an external metronome. As the frequency of the driver was increased we expected subjects to move from a 1:1 to a 1:2 mode of coordination after passing through a region of phase drift.

3.4.1 Participants

Nineteen right handed subjects (14 male and 5 female aged between 19 and 40 years.) took part in the experiment. Most of the participants were volunteers from an undergraduate class at Florida Atlantic University. All procedures were approved by the local Institutional Review Board for the protection of human subjects and participants signed consent forms before taking part in the experiment.

3.4.2 Apparatus

Participants placed the index finger of their dominant hand in a custom built manipulandum that restricted the motion of the metacarpophalangeal joint to a single plane (see Kelso

and Holt (1980) for a complete description). Unnecessary vertical and horizontal hand movements were restricted by a padding placed against the sides of the hand. An angle calibrated potentiometer situated above the axis of motion was used to measure the position of the index finger. Finger movement was sampled at 128Hz using an ODAU analog-digital converter connected to an OPTOTRAK 3010 system. The external metronome, consisting of a sequence of comfortably sounding beeps, each of 40ms duration, was sent to the ODAU unit and a pair of headphones.

3.4.3 Procedure

Handedness was first determined using the Edinburgh Handedness inventory. During the experiment the frequency of the metronome was increased from 2.5 Hz to 12 Hz in steps of 0.5 Hz. Thirty cycles of the metronome were presented at each frequency. At the end of each block of thirty cycles, the metronome frequency was stepped up by 0.5 Hz. A single trial consisted of twenty blocks (2.5 Hz to 12 Hz) that were presented continuously in order of increasing frequency. Each subject performed five trials. Subjects were asked to coincide peak flexion of the index finger of their dominant hand with the onset of the auditory metronome. Since we did not perform a similar task asking subjects to coincide peak extension with the metronome beat, we cannot test for phase transitions from extension to flexion with our present experimental design, although such effects are well known (Kelso, 1984). Prior to the start of the experiment participants were told, “if the initial pattern of movement changes, do not force yourself to go back to it. Just continue with the pattern you find most comfortable”.

3.4.4 Analysis

We first analyzed the metronome signal to find the onset of the beeps. This information was used to divide the movement data collected over a given trial into frequency plateaus. Each frequency plateau was coincident with a block of thirty metronome beeps. We noted that the finger movement over a given frequency plateau was almost sinusoidal. A discrete Fourier transform was performed and a single dominant frequency was obtained for each frequency plateau. The spectral power associated with other frequencies was relatively small. The

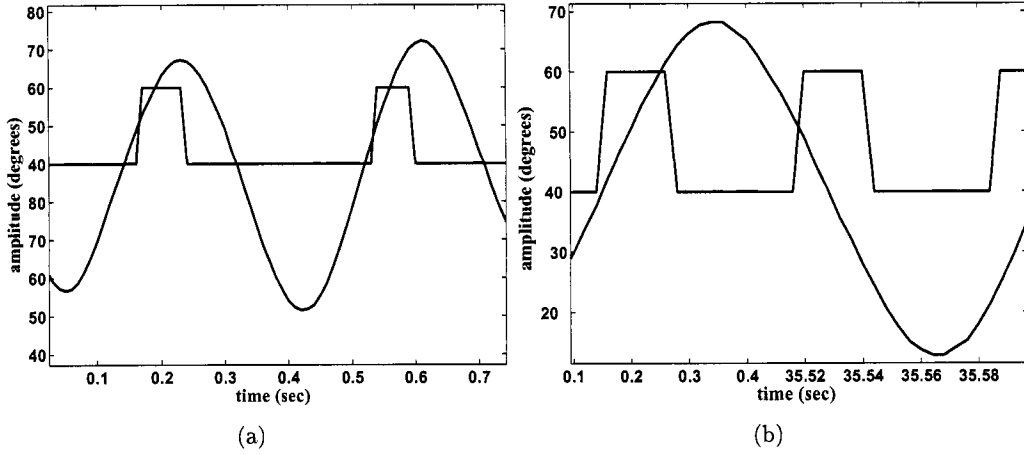


Figure 3.4: Classification of coordination patterns. The smooth curves are the trajectories of finger movement and the rectangular pulses represent the metronome. Amplitude is measured in units of the angular displacement (degrees). (a) An example of a pattern that was classified as 1:1 coordination and (b) Example of a pattern that was classified as 1:2 coordination. See text for details.

most prominent frequency for a given frequency plateau was chosen as the frequency of finger movement, ω , for that plateau.

In order to classify a particular frequency plateau as 1:1 or 1:2 the following classification scheme was used. Let t_P denote the time of occurrence of a local maximum of the finger movement and t_V a local minimum. The time of onset of a metronome beat is denoted by t_M . The cycles in which the patterns or $t_M \leq t_V \leq t_P \leq t_M$ occur were classified as 1:1 coordination cycles. The cycles in which the subjects performed the patterns $t_M \leq t_P \leq t_M \leq t_V$ or $t_M \leq t_V \leq t_M \leq t_P$ were classified as 1 : 2 coordination cycles. Examples of cycles that were classified as 1:1 or 1:2 are shown in Figures 3.4a and 3.4b respectively. If, during a particular frequency plateau consisting of thirty cycles, more than eighty percent of the cycles performed by the subject were 1:1 coordination cycles, then the frequency plateau was classified as a 1:1 coordination plateau. Similarly, the 1:2 coordination plateaus consisted of conditions in which more than eighty percent of the cycles were 1:2 coordination cycles. In our experiment, the task involved a range of metronome frequencies that were significantly higher than is traditionally used in finger movement studies since we were looking for subharmonic entrainment of the finger movement to the metronome. We sampled twenty frequency plateaus per trial. The sequence of metronome frequencies were presented in suc-

cession without a break in between frequency plateaus. Due to this experimental design, we were compelled to restrict the number of cycles per trial to thirty so that fatigue would not be a significant issue. However, using merely thirty cycles for higher frequencies implies that the length of each frequency plateau is small (A 5 second plateau for a metronome frequency of 6Hz and less for higher frequencies). Therefore, our frequency estimate has reduced accuracy with increasing metronome frequency. Hence, a weak ordinal criterion to identify mode locks (if 80 % of the cycles in a frequency plateau are 1:1 then the plateau is classified as such) had to be employed. Thus, it cannot be said that a stable mode locking was established.

3.5 Results

The results presented in this section will follow the steps in our theoretical analysis. In order to characterize the oscillator system and the coupling between its components we considered three aspects of the behavior.

1. The frequency of finger movement in relation to that of the metronome.
2. The phase of finger movement at the point of onset of the metronome.
3. The amplitude response for various driving frequencies.

3.5.1 Frequency

One of the primary goals of this experiment is to determine whether a consistent frequency relationship exists between the metronome and the movement of the subject's index finger. In Figures 3.5a-c , $\tilde{\Omega}$ is the frequency of the metronome and ω , the frequency of finger movement as defined above. The regions in which the subject maintained a fixed frequency ratio, $\frac{\tilde{\Omega}}{\omega}$, appear as horizontal plateaus in the figure.

Figures 3.5a-c presents the results from three subjects, representative of the broad range of behavioral patterns seen across all nineteen subjects. The ratio of the frequency of metronome to the frequency of finger movement, $\frac{\tilde{\Omega}}{\omega}$, was computed for each frequency

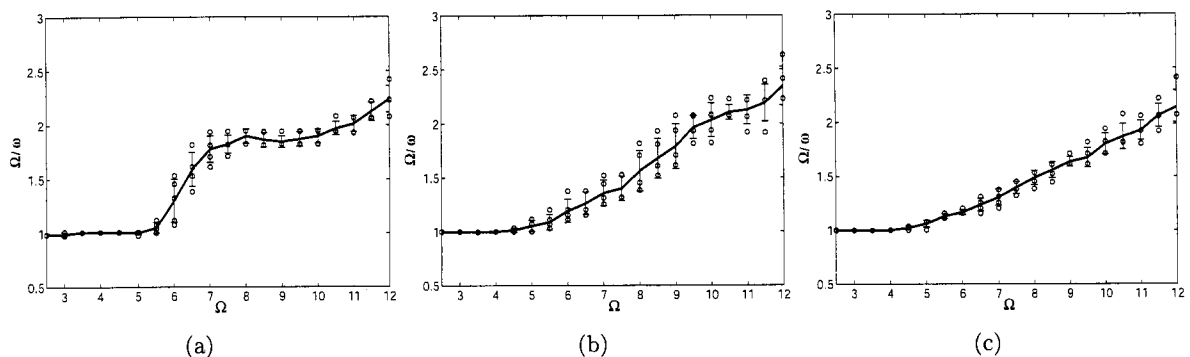


Figure 3.5: The circles show the frequency ratios for five trials at each frequency plateau, and the bars represent the standard deviation(see text for details). ω is the frequency of finger movement and $\tilde{\Omega}$ is the frequency at which the metronome is presented. The data shown in a, b and c are from three subjects and are representative of the patterns seen across all nineteen subjects.

plateau. The plotted ratio is the mean of five thirty cycle plateaus with the same frequency for a particular subject. The standard deviation across five trials for each subject is shown as error bars in Figure 3.5a. Most subjects show a distinct 1:1 mode of coordination followed by a region where the ratio $\frac{\tilde{\Omega}}{\omega}$ does not persist at a constant value. In the 1:2 regime the behavior varies continuously across subjects, ranging from a strong pattern of frequency locking (Figure 3.5a) to no apparent frequency locking (Figure 3.5b). Within the context of driven limit cycle oscillators, we consider three possibilities for the type of coupling. These are illustrated in Figure 3.6 which was generated by simulating a driven oscillator under different conditions of driving:

1. The coupling between the metronome and the finger is purely linear (Schöner and Kelso, 1988c). This is depicted by the dashed line in Figure 3.6. A linear coupling implies that the frequency of finger movement always coincides with the frequency of the driver. Hence, we would expect to see only one plateau in the figure, that of 1:1 coordination. This clearly violates the patterns of experimental results observed. Hence, a purely linear driving term is insufficient to explain our experimental findings.
2. The subjects were able to move their finger in a 1:1 pattern until a certain threshold frequency was reached beyond which they maintained a constant frequency. This

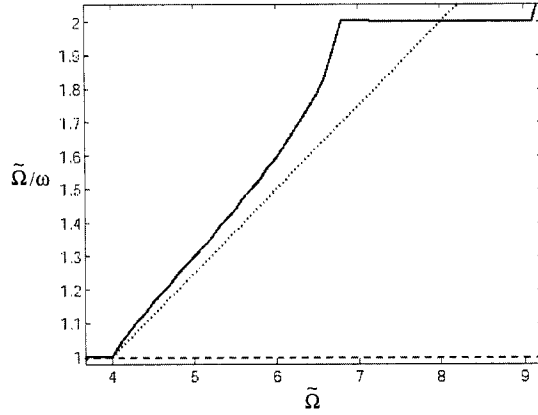


Figure 3.6: Plotted are the frequency of the driver $\tilde{\Omega}$ vs the ratio $\frac{\tilde{\Omega}}{\omega}$ for three possible coupling scenarios between stimulus and movement. The dashed line represents the profile due to purely linear driving obtained by setting $\epsilon_1 = 0.1$ and $\epsilon_2 = 0$. The solid line indicates parametric driving obtained by setting $\epsilon_1 = 0.1$ and $\epsilon_2 = 10$. The dotted line shows the behavior of a the oscillator which, above a maximum threshold frequency, exhibits a constant frequency of oscillation regardless of the frequency of the driver. The other parameter values are from Eq. 3.9 are $A = 1.5$; $B = 1$; $\gamma = 0.1$; $\omega = 4$

is depicted by the dotted line in Figure 3.6. Here, we assumed that the oscillator performs 1:1 coordination until a threshold frequency of 5Hz, beyond which a constant frequency is maintained throughout the range of frequencies of the driver. This compares with the behavior seen in Figure 3.5b where the subject performed 1:1 coordination until a frequency of 4.5Hz and then maintained an almost constant frequency (mean = 5.4Hz, standard deviation = 0.37) throughout the rest of the trial. However, this possibility alone does not account for the behavior seen in Figure 3.5a. Therefore, we must account for another possibility.

3. The coupling between the finger and the external driver is achieved using a parametric driving term. This is depicted in Figure 3.6 by the solid line. Here we see the existence of two plateaus corresponding to a 1:1 and a 1:2 mode of coordination. Compare this with the behavior seen in Figure 3.5a. Here the subject clearly shows the presence of two plateaus corresponding to the 1:1 and 1:2 modes of coordination.

The behavior seen across all nineteen subjects can be described in terms of varying contributions from the latter two forms of coupling and can be compared with the theoretical

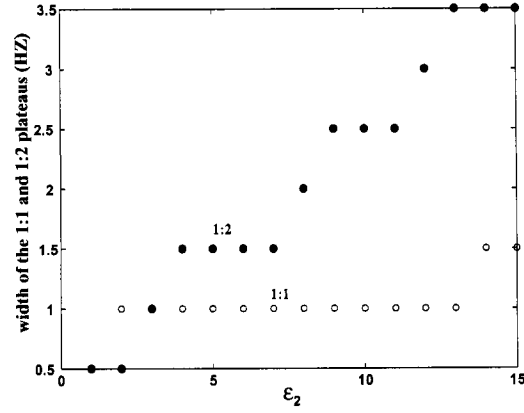


Figure 3.7: Plotted is the width of the 1:1 (empty circles) and 1:2 (filled circles) Arnold tongues as the strength of the parametric coupling term, ϵ_2 , is varied. See text for details.

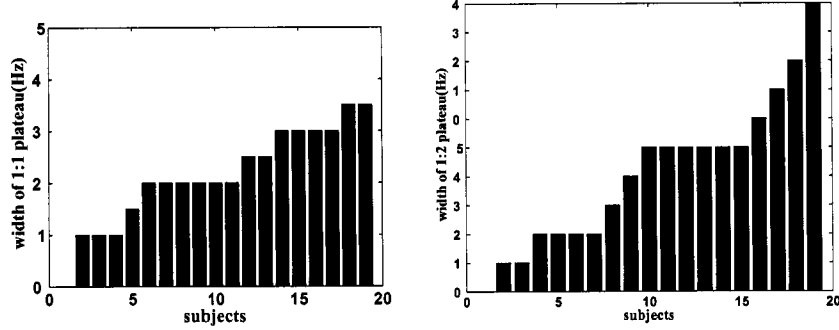


Figure 3.8: Each bar corresponds to a subject and the height of the bar corresponds to the width of 1:1 (a) or 1:2 (b) frequency conditions. See text for details.

model given by equation (3.9).

Our model predicts that the width of the Arnol'd Tongue structures increases as a function of the strength of the parametric (ϵ_2) driving term. This behavior is plotted as a function of ϵ_2 in Figure 3.7 for both plateaus, 1:1 and 1:2. In order to compute the width of each Arnol'd tongue for a particular driving frequency, $\tilde{\Omega}$, we first performed a discrete Fourier transform on the steady state time series obtained by numerically integrating Eq. 3.9. The most dominant frequency was then chosen as the frequency of the oscillator, ω . The width of the 1 : 1 Arnol'd tongue was determined by identifying the points within $|\omega - \tilde{\Omega}| < 0.1$. Similarly we computed the width of the 1 : 2 Arnol'd tongue by identifying the points within $|\omega - \frac{\tilde{\Omega}}{2}| < 0.1$.

In Figures 3.8a-b we describe the broad range of behavior seen across all nineteen subjects. Each bar in the figure corresponds to a single subject. The height of the bar represents the number of frequency plateaus in which the subject demonstrated a 1:1 (Figure 3.8a) or 1:2 (Figure 3.8b) mode of coordination. The procedure identifying the width of a synchronization plateau is described in section 3.4.4. In Figure 3.8a-b, the subjects are arranged in ascending order of the number of frequency plateaus in which each subject showed 1:1 (Figure 3.4a) or 1:2 (Figure 3.4b) modes of coordination. Here, 14 out of 19 subjects show significant plateaus (as defined in the section 3.4.4) of at least 1.5 Hz width for 1:1 coordination, whereas 11 out of 19 subjects show plateaus with at least 1.5 Hz width for 1:2 coordination. The variation of the width of the synchronization plateaus (Figures 3.8a-b) may indicate a variation in the coupling strength across subjects. The relationship between the coupling strength and the width of the plateaus for our model is shown in Figure 3.7.

3.5.2 Amplitude

In this section we examine two aspects of the dynamics, the amplitude response to a change in frequency and the effect of the metronome on the trajectory of finger movement.

3.5.2.1 Amplitude-Frequency relationship

As mentioned earlier, the movement is almost sinusoidal. Since the frequency variability within a given frequency plateau is small, we may use the spectral power associated with the most dominant frequency as a measure of the amplitude. Spectral power associated with the neighboring frequencies does not cause a qualitative change in the amplitude response function. Figure 3.9a shows the mean amplitude response to a change in frequency for all nineteen subjects. As the frequency of the metronome is increased the amplitude of finger movement decreases consistently across subjects. Different frequency-amplitude relationships, other than monotonically decreasing, have been observed over limited ranges of the frequency. Kugler and Turvey (1997) described an inverted U-shaped behavior for a bimanual wrist-pendulum system. The existence of a non-monotonic frequency-amplitude relationship has also been shown in single trials by (Beek et al., 2002). We found similar frequency-amplitude relationships for some trials although this was quite variable across

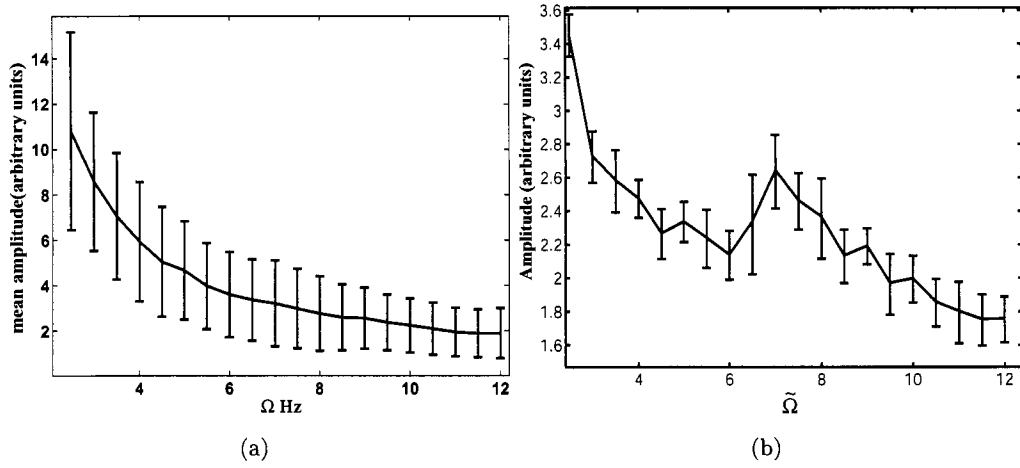


Figure 3.9: a) Plot of the mean amplitude as a function of the frequency of the metronome. The mean is obtained for each frequency plateau using data from nineteen subjects. The error bars show the standard deviation across nineteen subjects for all trials. b) The data are from one subject and averaged across five trials. Here, we can see the presence of two peaks at 2.5 Hz and 7 Hz. These features may correspond to the positions of the 1:1 and 1:2 Arnol'd tongues

subjects. In Figure 3.9b, for example, the maxima of the amplitude roughly coincide with the different frequency plateaus associated with the 1 : 1 and 1 : 2 Arnol'd tongues in Figure 3.5a. Both these figures are obtained using data from the same subject. Parametric excitation, as discussed in sections 3.3.1 and 3.3.2 provides a straightforward explanation of the amplitude response behavior observed in Figure 3.9b. However, the data averaged across subjects and trials shows only a monotonic decrease in amplitude (Figure 3.9a), and finer structure that might resolve the peaks in the region associated with the 1 : 2 Arnol'd tongue cannot be seen.

The experimental condition required subjects to coincide peak flexion with the metronome beat. In frequency regimes where the subjects performed a 1 : 1 coordination pattern, peak flexion coincided with the metronome. At the opposite reversal point, peak extension, the metronome was absent. Figure 3.10a was generated using the data from one subject performing in the 1 : 1 coordination mode at a frequency of 4Hz. Along the horizontal axis is plotted the position, and along the vertical axis, the velocity of the finger. The boxes show the regions corresponding to maximum flexion and extension in the phase plane. Compare the spread of the trajectories in the box marked 'flexion' with that in the box marked

ship. Earlier studies in bimanual coordination experiments have also shown the presence of this phenomenon, referred to in the literature as anchoring (Byblow et al., 1994). Fink et al. (2000) showed that the presence of a metronome at both peak flexion and extension reduces the propensity to switch from an anti-phase to an in-phase mode of coordination as compared to when the metronome is present only at extension. Thus the presence of specific environmental information affects not only local changes in the trajectory, but also induces changes in the global behavior of the system. The anchoring phenomenon demonstrated in our experimental results compares favorably with that seen in Figures 3.3a and 3.3b which was obtained by numerically simulating equation (3.9) in the presence of a Gaussian noise term.

3.5.3 Phase

The finger movement data for each subject were divided into frequency plateaus using the onset times of the auditory metronome as a reference. We calculated the continuous phase of the movement for each frequency plateau using a Hilbert transform. The points in the continuous phase trajectory that coincided with the onset times of the metronome were extracted. For each frequency plateau we obtained thirty such points corresponding to thirty cycles of the metronome. This method of evaluating the phase is equivalent to the commonly used point estimate of the phase for a given cycle, which is usually obtained using the time difference between a peak of finger movement and the onset of the nearest metronome peak. The trace in Figure 3.11 shows the standard deviation for each frequency plateau for each trial, averaged across all trials for a given subject. Variability across trials is denoted by the error bars. In Figure 3.11, the driving frequency, $\tilde{\Omega}$, is plotted along the x-axis and the ratio of the driving frequency to the finger movement frequency is shown along the y-axis. The figure is obtained exactly as in Figures 3.5a-c (see section 3.5.1). The two panels from Figure 3.11 are from the same subject and are shown together for comparison.

Figure 3.12a allows a further exploration of the causes of variability near the borders of Arnol'd tongues. Shown is a section of the time series of the continuous phase strobed using the onset time of the metronome. The segment of the time series shown corresponds to the edge of the 1 : 1 mode of coordination and the region of phase drift. In the initial part of

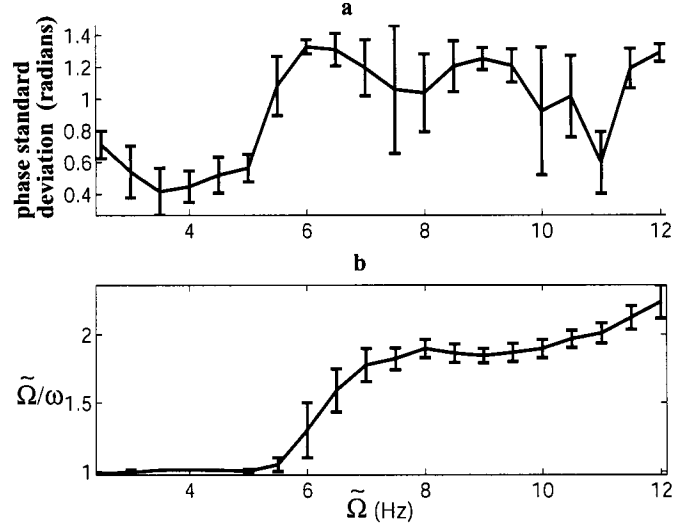


Figure 3.11: a) Standard deviation of the phase of finger movement as a function of the metronome frequency for one subject. Error bars represent the deviation of these values across five trials. b) A plot of the ratio $\frac{\tilde{\Omega}}{\omega}$ vs. Ω for the same subject. See text for details

the time series, the subject performs a 1 : 1 pattern with respect to the metronome, that is, the onset times of each metronome cycle coincide almost precisely with some phase of the movement cycle. The standard deviation of the phase in this region is small as seen in Figure 3.11. The data in Figures 3.11 and 3.12a-b are from the same subject. As the frequency of the metronome is increased, the subject departs from the 1 : 1 mode of coordination. (Between 40 and 55 sec in Figure 3.12a). This introduces a small frequency difference between the finger movement and the metronome. Each cycle of the metronome strobes the movement at a phase that systematically decreases over the successive cycles thus introducing a slope in the time series shown in Figure 3.12a (between 40 and 55 seconds). The rate of the phase drift is proportional to the difference in the frequency. In our experiment we are constrained to use a maximum of thirty cycles per frequency plateau. Over thirty cycles, the strobed phase is distributed between 0 and 2π . However, due to the small frequency difference, the entire $[0, 2\pi)$ interval does not get covered uniformly in thirty cycles. This is an artifact of the small number of points that are used to sample this interval. A consequence of this non-uniformity at the edge of an Arnol'd tongue is that, across trials, the variability of the phase standard deviation becomes large compared to the variability within an Arnol'd tongue. For

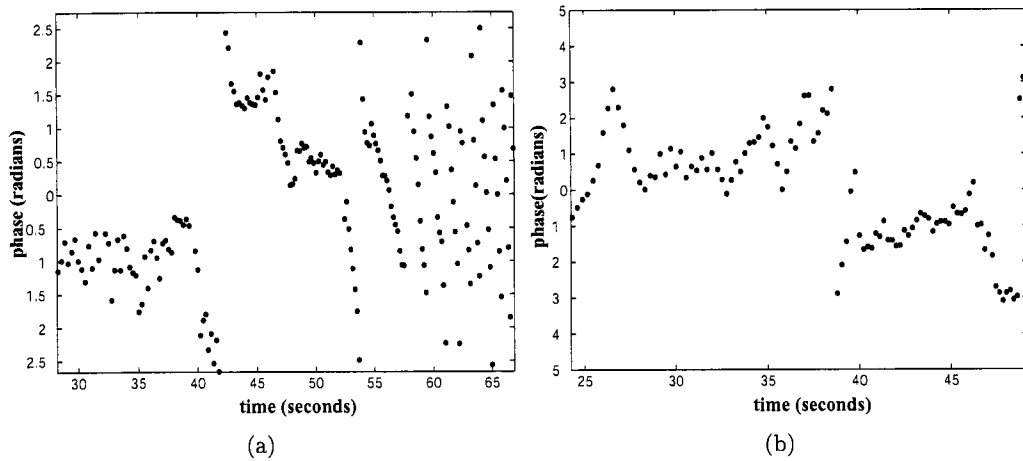


Figure 3.12: a) A plot of the phase of finger movement relative to the stimulus onset for one subject. A segment of the data on the border of the 1:1 mode of coordination is chosen. The behavior shows three distinct regions: a 1:1 coordination mode; a region where the drift is initiated; and a region between the 1:1 and 1:2 modes of coordination. Each point corresponds to the phase associated with one cycle of finger movement. b) Plot of the phase of finger movement at points of the movement coincident with a metronome beat obtained from a section of the experimental data for one subject. The data shows phase slippages at $t = 35$ sec and resetting at $t = 40$ sec

higher metronome frequencies in the region of drift, the points of the continuous phase that coincide with the onset of the metronome are distributed uniformly over the interval $[0, 2\pi)$. Therefore, in the drift region the standard deviation across thirty cycles of the metronome is large. However, due to the uniform distribution over the $[0, 2\pi)$ interval, the variability across trials is small. This leads to a maximum in the variability of the standard deviation across trials at the edge of an Arnol'd tongue. Similar maxima are observed at the edge of the 1 : 2 mode of coordination.

In a small region within the 1:1 coordination mode the subject showed phase drift and then resumed in a 1:1 mode of coordination. An example is shown in Figure 3.12b. This kind of behavior occurs in regions close to the edge of an Arnol'd tongue where a small perturbation, due to intrinsic or environmental noise, plunges the system into a region where the solutions are qualitatively different.

3.6 Mechanisms of phase transitions

One of the crucial features of the HKB model is its ability to characterize the observed phase transitions from anti-phase to in-phase mode of coordination that accompany scalar changes in cycling frequency (Kelso et al., 1981; Kelso, 1984). This phenomenon has also been established using more neurally based approaches (Nagashino and Kelso, 1992; Grossberg et al., 1997; Jirsa and Haken, 1997). Kelso et. al. (1990) have demonstrated a similar transition for the case of a single limb performing a rhythmic coordination task with a metronome. Subjects were asked to perform an off-the-beat rhythmic task, that is, peak flexion was coincident with a point between two metronome beats. As the frequency of movement was increased subjects spontaneously shifted from an off-the-beat to an on-the-beat pattern of coordination, analogous to a jump from an anti-phase to an in-phase coordination state seen in bimanual experiments. Research conducted using EEG and MEG (Kelso et al., 1992, 1999; Mayville et al., 2001) have consistently shown that a phase transition in behavior from an anti-phase to an in-phase mode of coordination is accompanied by changes in patterns of neural activity. In this section we will examine how such a transition can be effected and in the process postulate the existence of two types of transitions, a phase-mediated transition and an amplitude-mediated transition. Consider the equations of motion of the system in the plane spanned by the amplitude, r , and the phase, ϕ .

$$\dot{r} = f(r, \phi), \quad \dot{\phi} = g(r, \phi) \tag{3.36}$$

The above equations are referred to as the complete system. The complete system may operate on different time scales, for example, the time scale of change of the amplitude variable, r , may be radically different from that of the phase variable, ϕ . The presence of distinct types of transitions arise from a separation of time scales. The faster time scale can be adiabatically eliminated and the slower time scale then dictates the subsequent behavior of the system. For the case of a phase mediated transition the amplitude rapidly decays to its steady state value and the entire dynamics can then be described by the phase variable

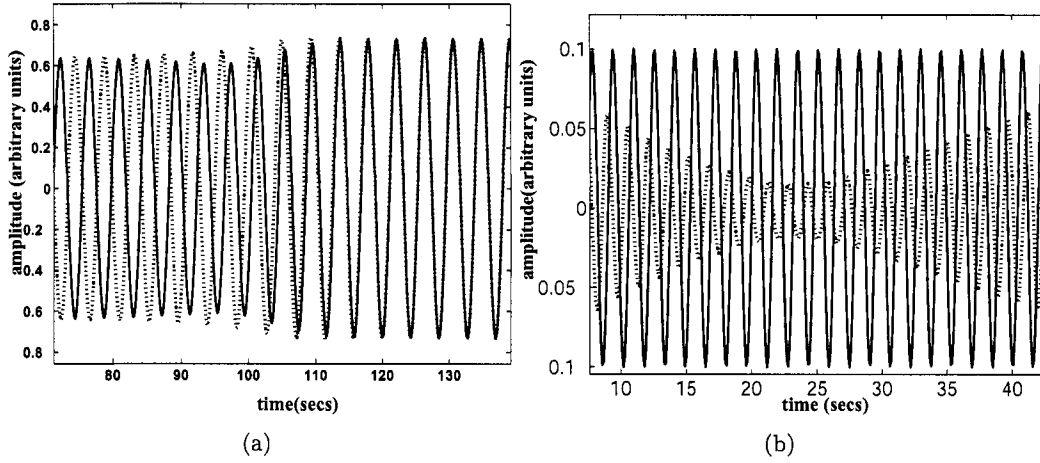


Figure 3.13: Two kind of phase transitions. a) Time series of a coupled oscillator system showing a phase mediated transition. The trajectories are obtained by integrating the HKB model (equations 3.1, 3.2 and 3.3). The transition takes place over a few cycles without a significant change in the amplitude of the oscillation. The solid and the dotted lines are the time series of each oscillator of the coupled oscillator system b) Time series of a parametrically driven oscillator given by equation (3.9) showing an amplitude mediated transition.

(Haken, 1983).

$$\dot{r} = f_1(\phi) \quad \dot{\phi} = g(f_1(\phi), \phi) \quad (3.37)$$

The HKB model is an example of a system that exhibits a phase mediated transition (Haken et al., 1985). Adiabatically eliminating the amplitude we arrive at the equations (3.4,3.5) for the dynamics. The numerically generated time series in Figure 3.13a shows that the transition takes place over a few cycles without any change in the amplitude.

For the case of an amplitude mediated transition the dynamics of the phase takes place at a time scale much faster than that of the amplitude. The steady state value of the phase can now be replaced in the amplitude equation to obtain the entire dynamics of the system.

$$\dot{r} = f(r, g_1(\phi)) \quad \dot{\phi} = g_1(r) \quad (3.38)$$

An example of an amplitude-mediated transition is seen in the equations for a parametrically paced version of the HKB oscillators. Transitions in coordination are usually described in polar coordinates, where the relative phase jumps through π and the amplitudes are positive. Alternatively, the phase can be held constant while the amplitude crosses over to a negative

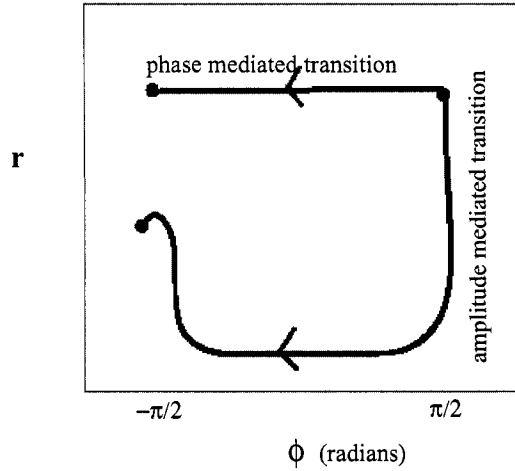


Figure 3.14: Two mechanisms of phase transitions in rhythmic coordination, an amplitude and a phase mediated transition in the plane spanned by the amplitude, r , and the phase, ϕ .

value. This description has been advanced by Jirsa et al. (2000). The time series shown in Figure 3.13a was obtained by numerically simulating equation (3.9). A phase transition is achieved through an amplitude mediated mechanism where the amplitude of the oscillation dies down and comes up again in the opposite phase.

Figure 3.14 illustrates the two different mechanisms for phase transitions. Along the horizontal axis is plotted the phase of the oscillator. The amplitude, r , is plotted along the vertical axis. In a phase mediated transition, described by equation (3.37), the amplitude remains constant while the phase performs a transition through π . In an amplitude mediated transition (3.38) the amplitude decreases and then grows, but in the opposite phase. With this general description of phase transitions in hand let us now examine the kind of transitions observed in our model. The equations of motion for the complete system are given by Equations (3.19,3.20). Here, we consider only the single metronome condition, that is, the frequency of finger movement coincides with that of the metronome. In the $\dot{\phi}$ equations we notice contributions due to linear and parametric driving terms. A parametric driving term

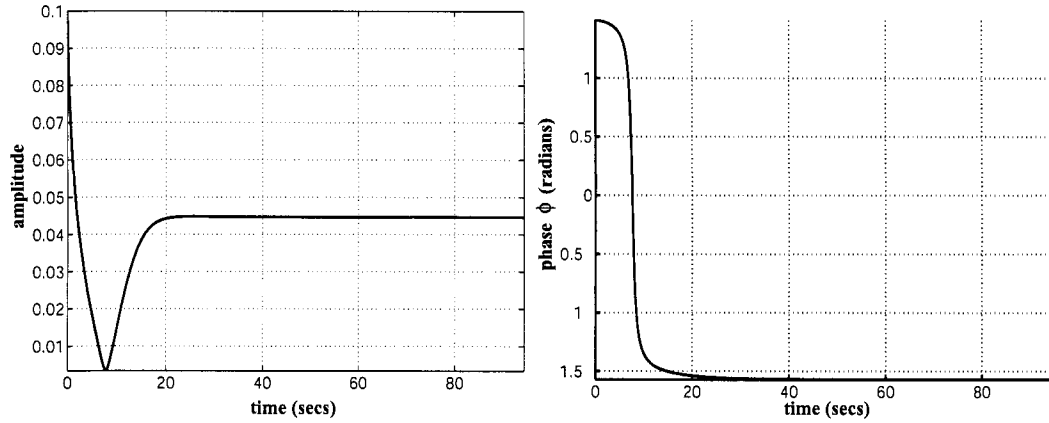


Figure 3.15: Plot of the time series of the amplitude, r , (a) and the phase, ϕ , (b) during an amplitude mediated transition obtained by simulating Equations 3.19 and 3.20. A drop in the amplitude is accompanied by a jump in the phase through an angle π . $A = 1.5$; $B = 1$; $\gamma = 1$; $\epsilon_1 = 0.1$; $\epsilon_2 = 3$; $\omega = 2$

leads to bistability, that is, there are two stable fixed points in the phase space of the system and depending on initial conditions the system tends towards one of the fixed points. Linear driving enforces monostability in the system by destabilizing one of the fixed points. The state of the system, monostability or bistability, is determined by the relative contributions of the two terms. From equation (3.20) we can see that the amplitude plays a crucial part in the transition. The contribution from the linear driving term depends inversely on the amplitude. As the amplitude of the oscillator approaches zero this contribution far outstrips that of the parametric driving term and the system becomes monostable. The local dynamics in the vicinity of a transition can be classified as an amplitude mediated transition. In Figures 3.15a and 3.15b we numerically integrated the equations (3.19) and (3.20) to obtain the time series of the amplitude and the phase of the complete system. Coincident with the drop in amplitude (Figure 3.15a), the time series of the phase (Figure 3.15b) shows a transition $\frac{\pi}{2}$ to $-\frac{\pi}{2}$.

Due to the presence of the $\frac{1}{r}$ term in the phase equation we cannot adiabatically eliminate the phase for the entire time course. However, in the neighborhood of a phase transition the trajectory in the r vs ϕ space does not show any change in the direction of the phase variable, i.e. $\dot{\phi} = 0$. Figures 3.16a-c describe the dynamics of the system in the $r - \phi$ plane. Consider the set of points with $\dot{r} = 0$. These are called the r nullcline. The set of points with

$\dot{\phi} = 0$ are called the ϕ nullcline. The arrows shown in Figures 3.16a-c point in the direction of the flow of the vector field given by equations (3.19,3.20). At points on the r nullcline the r component of the flow is zero, that is, the arrows point in the horizontal direction. At points on the ϕ nullcline the ϕ component of the flow is zero, that is, the arrows point in the vertical direction. The intersections of the r and the ϕ nullclines are the fixed points of the system.

Figures 3.16a-c show the nullclines and the flow directions in the $r - \phi$ plane for different values of the strength of the linear driver, ϵ_1 . As mentioned earlier, the relative contributions of linear, ϵ_1 , and parametric, ϵ_2 , driving terms determine monostability or bistability of the system. In Figure 3.16a $\epsilon_1 = 0.05$. Here, the contribution due to the linear driving term is negligible and the system shows the presence of three fixed points, two stable and one unstable. As ϵ_1 is increased to 0.1 in Figure 3.16c, one of the stable fixed points and the unstable fixed point coalesce and disappear and all the trajectories converge to the remaining stable fixed point, that is, the system becomes monostable. Preliminary data from (DeGuzman et al., 2002) provide some confirmation of an amplitude mediated transition mechanism for the case of bimanual coordination. The experiment compared data from two experimental paradigms, that is, coordination of cyclical finger movements with and without a metronome. DeGuzman et al. (2002) report that in the presence of an external metronome subjects perform a transition from anti-phase to in-phase coordination states via an amplitude-mediated mechanism in 70% of all trials examined. In contrast, in the self-paced case, that is, coordination in the absence of a metronome, subjects showed phase-mediated transitions in about 80% of the trials examined. The self-paced case, can be clearly described using the original HKB equations. However, in order to describe the results for the metronome-paced case, one would require the presence of a linear and a parametric driving term.

3.7 Summary and conclusions

The chief motivation behind this study was to examine the nature of action-perception coupling in the coordination dynamics of movement. Rhythmic movements, driven by a

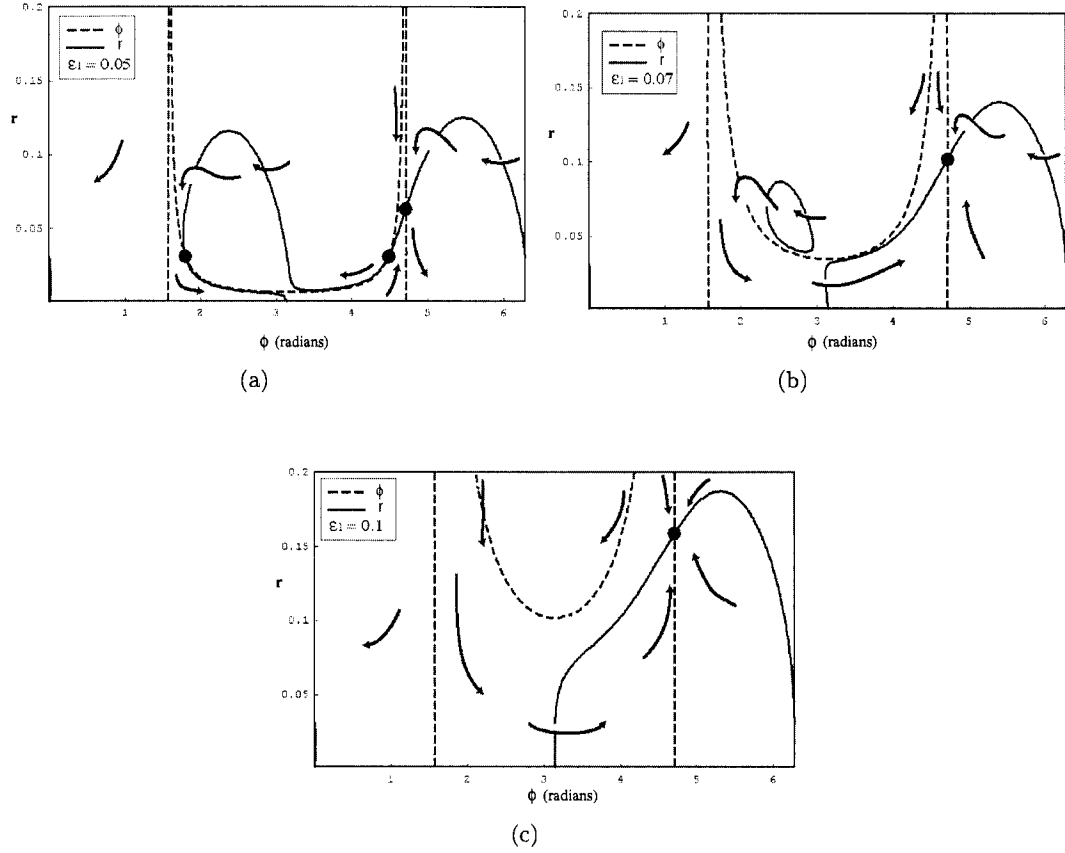


Figure 3.16: Plots of the amplitude, r , and the phase, ϕ , as a function of increasing the strength of the linear coupling, ϵ_1 showing the mechanism of an amplitude mediated transition for the equations (3.19,3.20). The solid dots represent the fixed points of the system. The arrows show the approximate direction of the flow in the space spanned by the amplitude, r , and the phase, ϕ . See text for details.

periodic stimulus, present a paradigm that is ubiquitous in nature and has the additional advantage of being accessible in terms of well studied ideas of oscillator theory. Numerous experimental and theoretical studies have demonstrated the appositeness of this approach. We employed a limit cycle oscillator, based on the Haken-Kelso-Bunz model, to describe the intrinsic dynamics of the system. The external metronome was operationalized by a periodic external stimulus. Our central hypothesis was that the coupling between the external metronome and the intrinsic dynamics of the system cannot be described by a purely linear driver. Therefore, a minimal description of the system must use a parametric driving term. Here, we described an experimental strategy to verify this hypothesis. Parametrically driven limit cycle oscillators are characterized by the presence of Arnol'd tongue structures (as shown in Figure 3.1). The experiment consisted of driving the system through different regions of the parameter space in order to extract signatures of the Arnol'd tongue structures using frequency, amplitude and phase measures. The synchronization behavior exhibited by subjects clearly showed the presence of 1:1 and 1:2 coordination regimes. Signatures of the 1:1 and 1:2 Arnol'd tongue structures were also seen in the reduced standard deviation of the phase associated with these regions (section 3.5.3). Experimentally observed anchoring phenomena were shown to be a consequence of parametrically driving the limit cycle oscillator. The model presented accounts for bistability, and a transition from bistability to monostability as observed in unimanual coordination experiments. The main new feature of our theoretical model is the existence of two mechanisms via which a phase transition can be achieved, an amplitude mediated transition and a phase mediated transition. The transitions follow mutually orthogonal paths to a steady state. A possible quantification of the degree of phase or amplitude mediation can be described by examining the movement trajectories in the space spanned by the amplitude and phase of finger movement. Before a transition, the fluctuations in the direction of the slower variable should be significantly lower than that of the fast variable. This prediction is now open to further experimental test.

Chapter 4

Asymmetries in temporal multisensory perception.

4.1 Introduction

Our experience of the external environment is influenced by the constant influx of sensory information into different sensory pathways. A task of the brain is to combine these seemingly parallel streams of information into a coherent picture of the world. The reception, transduction, and processing of sensory specific information is characterized by different temporal and spatial acuities (Andersen et al., 2004). Although in our daily experience we perceive information from different modalities coherently, that is, in spatio-temporal register, there is evidence to suggest that different modalities are perceived with different temporal delays. In addition, we know that multisensory integration is mediated by a widely distributed network of brain regions (Calvert, 2001). The communication between these regions must necessarily be subject to delays. In the present study we will be primarily concerned with the temporal features of multisensory integration. How is information from different sensory modalities combined inspite of the existence of temporal asymmetries in the perception of individual modalities and delays in the propagation of information across distributed brain regions? An ideal testing ground for the temporal features of multisensory integration is related to the perception of simultaneity. The objects and events in our environment generate several

simultaneous physical energies. A stone hitting the ground generates visual and auditory cues in addition to vibrations of the ground that produce tactile cues Meredith (2002). These physical energies do not interfere with each other as they impinge on the respective sensory receptors. However, we associate these concurrent physical effects as emanating from the same object or event. The simultaneity of these different physical energies provides a powerful cue to determine whether they are associated with the same event or object. The key issue in the context of audiovisual stimuli may be expressed as follows: Light travels at a speed approximately six orders of magnitude faster than sound. However, the visual system is sluggish relative to the auditory system in part due to slow molecular changes in the retinal photopigment that are induced by light. These changes are in the microsecond range. The temporal differences in propagation and conduction of auditory and visual stimuli cancel each other at a distance of 10 meters, the so called ‘horizon of simultaneity’. Most physical stimuli that occur simultaneously do not occur at this distance. Nevertheless, the stimuli are perceived as simultaneous (Spence and Squire, 2003).

A number of indicators of the difference in latencies between the auditory and visual systems is provided by reaction time experiments. Kling and Riggs (1971) note that the simple reaction time to a strong visual stimulus reaches a minimum of 150ms. In comparison strong auditory and cutaneous stimuli yield reaction times that are 30-40ms shorter. Both auditory and cutaneous stimuli are not constrained by lengthy intermediate processing as seen in the retina. Afferent responses to touch, pressure and vibration are conducted by large myelinated fibers in the peripheral nerves. Further, the latency of the first signature of activity in the visual cortex, identified by the early C1 wave in event related potential (ERP) studies is noted to be approximately 46ms (Foxe and Simpson, 2002). In comparison, ERP studies using auditory stimulation demonstrate auditory cortex activity 15-20ms post-stimulus. Temporal differences in conduction and processing of environmental inputs from different modalities clearly exist. In a multisensory context, for example with an auditory-visual or a tactile-auditory stimulus, these temporal asymmetries may lead to asymmetries in perception. There exists a vast literature detailing temporal asymmetries in the perception of auditory-visual multisensory stimuli in particular. The general findings, though divergent in their specifics, consistently demonstrate asymmetries in the perception of auditory-visual

simultaneity, with people more likely to perceive an auditory-visual stimulus as simultaneous if the auditory stimulus precedes the visual stimulus (Spence and Squire, 2003; Levitin et al., 2000). The estimates are highly variable across subjects. The window of time and the nature of this window over which the two stimuli are perceived as simultaneous is a matter of much controversy (see Sugita and Suzuki (2003) and Arnold et al. (2005) for differing points of view regarding the window of temporal multisensory integration). Differences in the temporal latencies of perception are not restricted to different sensory modalities but may also occur for different features within a particular sensory modality. For instance, Moutoussis and Zeki (1997) showed in a series of experiments that the perception of color precedes that of motion.

In order for different sensory modalities to interact, Stein and Meredith (1993) argue that the inputs from these modalities must converge into neurons that respond to stimulation from multiple modalities. Based on extensive anatomical and electrophysiological evidence (Meredith, 2002), they identify the superior colliculus as a major ‘convergence zone’. Neurons in the superior colliculus are known to be responsive to stimulation from multiple modalities. Spatially and temporally congruent stimulation from multiple sensory modalities generates neuronal firing that far exceeds what may be observed by summing the responses obtained by stimulation in each modality. Neuroanatomical studies in primates have identified numerous areas where afferents from different senses converge. At the cortical level, these so called multisensory convergence zones include the superior temporal sulcus (STS), intraparietal sulcus (IPS), posterior insula and frontal regions including premotor, prefrontal and anterior cingulate (AC) (Jones and Powell, 1970; Seltzer and Pandya, 1989; Mesulam and Mufson, 1982). Subcortical anatomical convergence zones include the claustrum, the superior colliculus and the hippocampus (Mesulam and Mufson, 1982; Pearson et al., 1982; Mufson and Mesulam, 1984). Electrophysiological studies have shown that many of these areas contain cells responsive to stimulation in more than one modality (Desimone and Cross, 1979; Vaadia et al., 1986). In particular, the cells in the superior colliculus have been investigated in detail with respect to crossmodal integration (Stein and Meredith, 1993). Evidence from brain imaging studies for auditory-visual integration sites has been found in the superior temporal sulcus using speech and the superior colliculus (Calvert et al.,

2001) and intraparietal sulcus (IPS) (Bushara et al., 2001) using non-speech stimuli. Giard and Peronnet (1999) found interaction effects in visual and auditory cortices resembling the latency, polarity and topography of the sensory-specific visual and auditory responses to the unimodal stimuli (visual P200 and auditory N100, respectively), suggesting multisensory enhancement of, but not necessarily integration in the primary sensory cortices. Banati et al. (2000) and Macaluso et al. (2000) used visual-tactile stimulation and PET to identify the STS, IPS and frontal areas as potential multisensory integration sites. Given the above neuroanatomical, electrophysiological and human imaging literature a pattern emerges indicating that a network of STS, IPS, frontal cortex, insula claustrum and superior colliculus act as multisensory convergence zones. There are also feedback connections to and between primary sensory cortices (Rockland and Ojima, 2001, 2003; Falchier et al., 2002), but their involvement in the actual multisensory integration process is an open question.

The evidence presented above suggests that multisensory integration is mediated by a widely distributed network of brain areas. The communication between these brain areas must necessarily be subject to temporal delays in propagation. Typical times for the propagation of an action potential along myelinated axons are approximately 6-9m/s. Time delays in propagation across adjacent gyri separated by distances of approximately 1cm are of the order of 1ms. However, the delays along long axons (20cm) connecting the frontal and occipital regions are of the order of 20-30 ms which is a significant delay in communication (Nunez, 1995). In unmyelinated axons the speed of signal conduction is approximately 1m/s (Kandel et al., 1991). A distance of a millimeter between neurons would lead to the conduction time delay of approximately 4% the period of an oscillation assuming 40Hz oscillations which seems to be an important time scale in feature binding (Tallon-Baudry et al., 1996). Therefore, in trying to understand the mechanisms of multisensory integration, which require communication across distributed networks, one must account for the temporal scales associated with the individual modalities and the temporal delays associated with the communication across distinct areas of the brain involved in the process of multisensory integration.

The goal of the present research is to understand the mechanisms of multisensory integration given the existence of multiple time scales of modality specific information processing

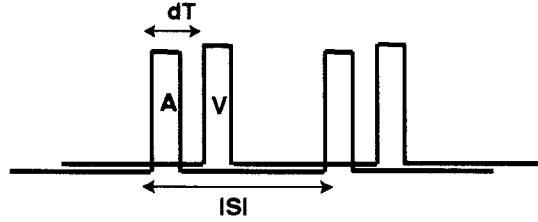


Figure 4.1: Shown is an example of the stimulus sequence for A (auditory) and V (visual). The parameters of the experiment are dT , the inter-onset interval between two stimuli of a multisensory pair and ISI , the time difference between two consecutively presented stimuli of a particular modality. Similar stimulus sequences were used for Haptic-Visual and Haptic-Auditory stimulation.

and temporal delays that exist in the communication between brain regions. In this paper we examine these issues along two directions. First, we report the results of a behavioral experiment using pairs of rhythmically presented multisensory stimuli. Our experiments are specifically designed to probe the temporal asymmetries and differential time scales associated with different sensory modalities and their interaction. Second, we examine our results in the context of coupled dynamical systems with different intrinsic time scales and temporal delays in communication that account for the experimental data and the results of past explorations into the nature of temporal multisensory integration.

A generalized synchronization paradigm was developed in order to test temporal features of multisensory integration parametrically. The stimulus sequence consisted of pairs of rhythmically presented stimuli. Each stimulus of the pair belonged to a different sensory modality (auditory, visual or tactile). The parameters varied during the experiment are the interval between the onsets of two, dT , and the inter-stimulus interval (ISI) (See Figure 4.1).

This design leads to a two dimensional parameter space. Each trial of the experiment samples one point in the dT - ISI parameter space. The experiment was repeated for all stimulus pairs (auditory-visual, tactile-visual, and auditory-tactile). Subjects were required to report their perception of the stimulus sequence according to a forced choice paradigm. The reported perception of the subject for every sampled point in this space was used to partition the parameter space, the symmetries of which, enable us to characterize some of the temporal features of multisensory integration. In our experiment and subsequent theoretical model we address the following question: How do differences in intrinsic time scales and

temporal delays affect the multisensory perception? Our behavioral experiment employs stimuli from three different sensory modalities (auditory, visual and tactile). We hypothesize that asymmetries in the intrinsic time scales associated with the different modalities must lead to asymmetries in perception.

4.1.1 Subjects

18 subjects with ages ranging from 24 to 40 participated in the experiment. All subjects reported normal hearing and corrected to normal vision. The Institutional Review Board of Florida Atlantic University approved all experimental procedures and subjects were required to provide written consent prior to the experiment.

4.1.2 Stimuli

Auditory: The auditory stimulus consisted of a sequence of beeps presented through a speaker placed in front of the subject at a distance of 2.3 meters. Each beep was a pure tone (440Hz) lasting 30ms with an intensity of 64dB.

Visual: The visual stimulus consisted of a series of flashes from a red light emitting diode placed at the same location as the speaker at the center of the visual field. The diameter of the light emitting diode was 10mm and produced a stimulus with a luminous intensity of 300 millicandelas.

Haptic: The haptic stimulus consisted of a series of bursts of a 300Hz sine wave presented using a MiniBuzzer placed on the subjects right index finger. The MiniBuzzer operated in a voltage range between 1.5V and 3.0V. We used a stimulus of amplitude 2V. The maximum response of the instrument was in the range of 300Hz to 500Hz.

4.1.3 Experimental conditions

The experiment was conducted over three sessions. Each session was held on a different day. During a given session a single pair of multisensory stimuli (Auditory-Visual, Tactile-Auditory, Visual-Tactile) was presented. The parameters of the experiment were the inter-onset interval between each stimulus of the pair (dT) and the frequency of stimulus

presentation ($f = 1/\text{ISI}$). A session consisted of 69 trials, each trial lasting 45 seconds. During each trial, a stimulus sequence with a specific dT and ISI was presented. The frequency ($f = 1/\text{ISI}$) was varied randomly from 0.5Hz to 3.5Hz in steps of 0.5HZ. The inter-onset interval within the stimulus pair was varied between 0ms and 200ms in steps of 50ms.

4.1.4 Instructions

Throughout this section, unless stated otherwise, we will refer to a specific stimulus pair (Auditory-Visual) for reasons of convenience. The set of instructions and experimental conditions was replicated for the remaining two stimulus pairs, Haptic-Visual and Haptic-Auditory.

Subjects were instructed to perceive the multisensory stimuli as **synchronized** events, maintaining a constant level of attention toward either stimulus modality. Subjects were told that it would not be possible to perceive the stimuli as synchronized events across all trials. The instruction specifically addressed the issue of different strategies that subjects may employ when presented with a stimulus sequence. We elaborate on this point later in this section. Subjects were specifically instructed that their task was not to guess the correct temporal precedence of the physical stimulus. It was to report **their perception** of the stimulus sequence. At the end of each trial subjects were required to report their perception of the stimulus using one of four options.

1. The auditory stimulus precedes the visual stimulus. (AV)
2. The visual stimulus precedes the auditory stimulus. (VA)
3. The two stimuli are simultaneous. (S)
4. Can't tell. (D)

The last option, 'Can't tell' is associated with a unique multisensory illusion. In our pilot studies we noticed that for certain points in the dT-ISI parameter space the perception of the stimulus sequence did not remain constant. The temporal order of the stimuli changed across the length of a 45 second trial akin to the drifting phase between two oscillators. Subjects were asked to report option 4 if their perception of the stimulus changed any

time during the trial. Typical reports of subjects on their perception include “the stimuli drifted relative to each other”, “the auditory and visual stimuli appeared as two different objects”. This is quite similar to a phenomenon in auditory perception, first reported by van Noorden (1975), known as streaming. He used a set of tones, differing in frequency and inter-onset interval, to elicit the perception of multiple streams. In some regions of our experimental parameter space we observe a phenomenon that may be considered as a multisensory analog of auditory streaming. However, it is important to note here that drift and streaming are two different percepts, both of which are contained within the domain of the ‘Can’t tell’ option. In his experiments on auditory streaming, Van Noorden reported the existence of three qualitatively different perceptual domains (van Noorden, 1975). A variation of stimulus parameters drove the perception of the stimulus from one domain to another. An important feature of this change of perception is related to the instructions Van Noorden gave his subjects. There exist some regions of the parameter space where subjects are able to perceive two different percepts depending on the instructions given, a state we will refer to as multistable. Here it is important to indicate the relevance of this observation to the instructions used in our experiment. Subjects may perceive the stimulus using some particular strategy, for example, attending to the auditory stimulus and not the visual stimulus. This would introduce additional variability in the results due to different strategies that produce different percepts, especially in multistable states. Our instructions, “try to perceive the stimuli as synchronous” addressed this specific issue. In order to account for differences that may arise due to different strategies employed by subjects in viewing the stimulus we included an instruction to try and perceive the multisensory stimuli as simultaneous or synchronous. This instruction is particularly pertinent in the neighborhood of parameter values where subjects switch from one percept to another. This region is bistable, in that there is more than one possible stable percept. We tested this bistability in later experiments using fMRI (Dhamala et. al. 2005) which demonstrate conclusively that each percept is characterized by a different set of brain networks though the stimulus presented was exactly the same. The percept of asynchrony was accompanied with the activation of the inferior parietal lobule, which was absent when the stimuli were perceived as synchronous. This finding is consistent with previous literature on the synchronous vs

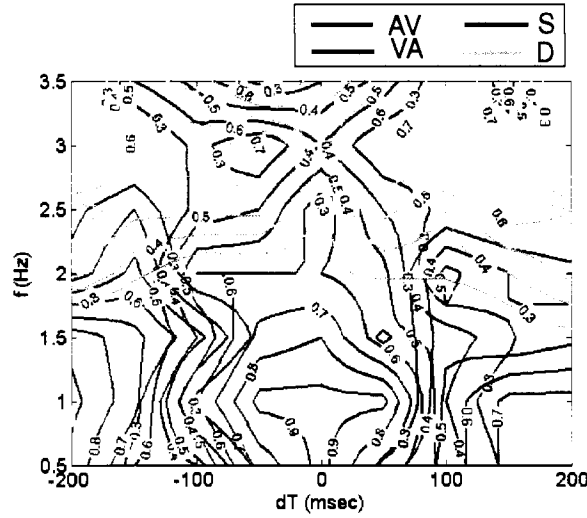
asynchronous multisensory perception (Bushara et al., 2001).

4.2 Experimental results.

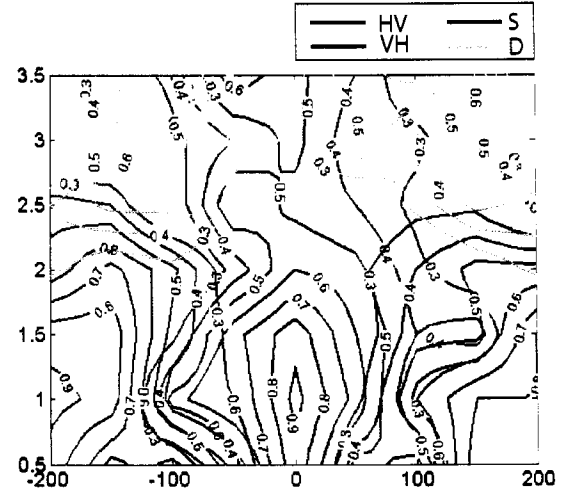
Various regions of the perceptual partitioning in the parameter space by the correspond to a particular percept. For example AV indicates that the perception of the Auditory stimulus preceded that of the Visual stimulus. S is the region where subjects perceive the stimuli as Simultaneous while D or ‘Drift’ indicates the region where subjects reported the ‘Can’t tell’ option.

Our primary interests lie in characterizing and interpreting the structure and symmetries of the regions in the dT-ISI parameter space that correspond to different perceptions of the stimulus sequence. An overview of the results for the three different pairs of stimuli is shown in Figures 4.2a-c. The inter-onset interval (dT) is plotted along the x-axis and the different frequencies of stimulation ($1/ISI$) are plotted along the y-axis. In Figure 4.2a, negative values of dT indicate that the auditory stimulus preceded the visual stimulus. In Figure 4.2b negative values of dT imply that the tactile stimulus preceded the visual stimulus while in Figure 4.2c it indicates that the tactile stimulus preceded the auditory stimulus. The maps in the Figures show the probability that a subjects reported a particular percept for a given dT-ISI point. The probability is given by the ratio of the number of subjects who reported a particular percept at a point in the two dimensional parameter space to the total number of subjects. The maps shown if the Figure are obtained using data from all subjects. The different perceptual states are shown in different colors according to the legend in the Figures. The contour lines for each percept connect points of equal probability in the dT-ISI space. The contour lines are truncated for probabilities below 0.3. In what follows we will determine the width and the asymmetries of the region of simultaneity for each stimulus pair. The other regions are adjacent to the region of simultaneity and any asymmetries in S would also be seen in the other regions.

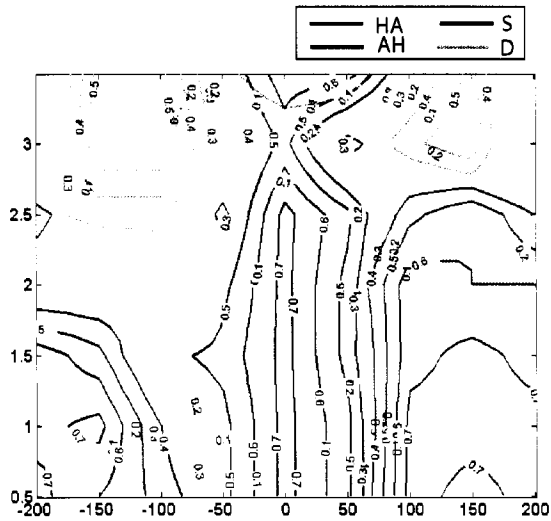
These maps present a broad overview of the perceptual partitioning of the parameter space. The most notable feature of these maps is the qualitative similarity between them. Though differences exist in the widths and the symmetries of various regions, the general



(a)



(b)



(c)

Figure 4.2: a) Shown are the probabilities that subjects perceive AV, VA, S or drift in different parts of the parameter space. The contour lines connect points where the probability that subjects report a given percept are equal. b) probabilities that a given subject perceives HV, VH, S or drift c) probabilities that a given subject perceives HA, AH, S or drift. The results shown are obtained using data from all subjects.

location of the simultaneous, sequential and drift percepts are preserved. Figure 4.3 shows the width of the region of simultaneity as a function of increasing frequency. The width of S for a particular subject is defined as the sum of the number of simultaneous responses reported by the subject for a particular frequency value. The width is measured in the units of dT (ms). The plot in Figure 4.3 represents the mean width of this region across subjects along with the standard error of means, with each color being associated with a particular pair of stimuli. Here it is notable that the width shows a mean decrease as a function of frequency except for a frequency of 3.5Hz. At this frequency subjects reported the perception of simultaneity for some non-zero values of dT. These points may be seen in the contour map plots as an asymmetrically located region of simultaneity near 3.5Hz. Such segregated regions of synchrony indicate the presence of inherent nonlinearities in the system. The magnitude of widths of the S region, measured in ms, for the three different stimulus pairs shows a trend in its ordering at low frequencies with the width of the S region being largest for auditory-visual stimuli and smallest for haptic-auditory stimuli.

We compared pairs of stimulus sequences, for example auditory-visual vs haptic-auditory, to determine whether the width of the S region was significantly different across all subjects as a function of the presentation frequency. The most significant difference between the width of the S region for auditory-visual stimulation compared to haptic-auditory stimulation is at 1Hz ($p < 0.05$). The only significant difference is confined to a comparison between auditory-visual and haptic-auditory stimulation. However, no significant differences were seen in the comparison of the widths of haptic-visual stimulation with that of haptic-auditory stimulation. We know that processing of visual stimuli are relatively slow due to slow chemical changes induced by light in the retina. Auditory stimuli, on the other hand, are propagated much faster. In the following section, based on our model, we will propose that the slow speed of visual processing is one of the reasons for the large width of the S region for auditory visual stimulation. In comparison, haptic and auditory processing are relatively fast which results in a window of simultaneity that is not as wide as for auditory-visual stimulation. Another source of the difference in widths may be related to the physical location of the individual stimulus sources. The auditory and the visual stimulus sources are presented from the same spatial location, while the haptic and the auditory stimuli are

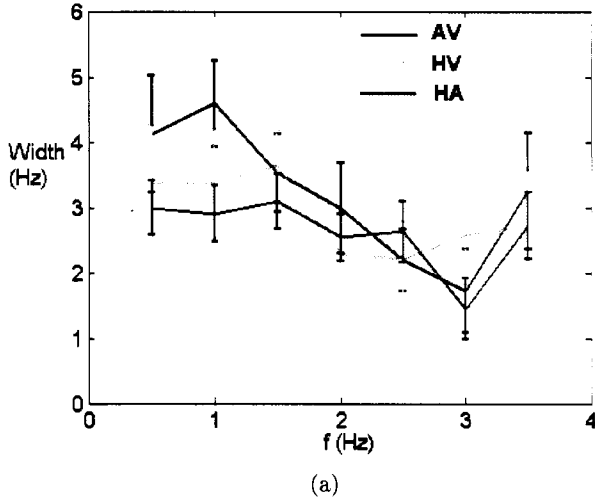


Figure 4.3: The mean width of the S region across subjects for different pairs of stimuli as a function of presentation frequency. The asterisk indicates the widths that are significantly different at a significance level of $p < 0.05$.

presented from different locations. A difference in spatial location may influence the ability of the subject to perceive the stimuli as simultaneous therefore affecting the width of the S region.

Further, we investigate the asymmetry of the region of simultaneity. A measure of the asymmetry of S may be obtained by calculating the number of S responses on either side of the $dT=0$ line. Figures 4.4a-c show a comparison of the number of S responses to the left (black line) and right (blue line) of the $dT = 0$ line as a function of the frequency. The most significant asymmetries ($p < 0.05$) are marked by an asterisk in the figures. The asymmetry in the perception of the Auditory-Visual stimulus sequence is tilted toward the perception of AV. Further, we observe asymmetries associated with Haptic-Visual and Haptic-Auditory stimulation. In addition to the temporal aspects of multisensory processing the perception of Haptic-Visual and Haptic-Auditory stimulus sequences is biased due to the different spatial locations of the respective sources. The remaining regions where subjects perceive AV, VA, or drift are adjacent to the region S. Therefore any asymmetries in the region S will also be seen in the remaining regions. In the following model we will propose that the asymmetries observed in the perceptual partitioning is a function of the asymmetries in the intrinsic times

scales associated with the processing and communication between the different modalities.

4.3 Model

Our data suggest that the perceptual partitioning of the parameter space is qualitatively independent of the different modalities involved. The differences that exist are in the symmetries of the different regions and their widths. The broad features, for example, an S region flanked by AV and VA in addition to a drift regime for higher frequencies, are replicated across stimulus pairs suggesting that these features may be a function of the coupling and the time scales associated with the different modalities. Qualitative changes in the coordinative behavior and perception have often been associated with corresponding qualitative changes in brain dynamics (Kelso et al., 1998; Fuchs et al., 1992). In chapter 1 (see also Bressler and Kelso (2001)), we argue that the functional units of processing are not single neurons, but the coordinated activity of ensembles of neurons. Assisi et al. (2005b) (chapter 2) demonstrated that the behavior of globally coupled ensembles of dynamical systems is approximated by a low dimensional system of equations. The Fitzhugh-Nagumo neuron (Fitzhugh, 1961; Nagumo et al., 1962) is a paradigmatic example of an excitable system that is of particular relevance to our modelling approach. An excitable system is characterized by the existence of a threshold. A supra-threshold perturbation to the system leads it along a large amplitude trajectory in phase space before settling down to a fixed point. Consider the example of globally coupled Fitzhugh-Nagumo neurons. The collective behavior of this system, which includes synchrony, phase clustering and oscillator death may be approximated by using a system of two coupled excitable elements Assisi et al. (2005b). For the case of synchronous behavior the mean activity of the ensemble of neurons may be approximated by a single Fitzhugh-Nagumo neuron. The dynamics of coupled excitable systems thus provides a useful starting point to discuss the dynamics of perception. For example, Morein-Zemir et al. (2003), in a series of behavioral experiments demonstrated that the perception of audio-visual simultaneity is related to the convergence of the perception of the visual stimulus toward the temporal location of a subsequently presented auditory stimulus. This convergence occurs only if the auditory stimulus is presented within a certain

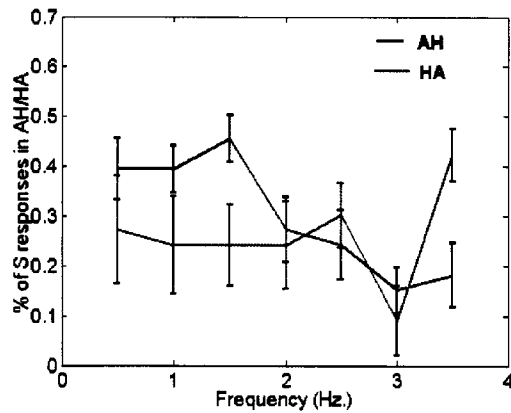
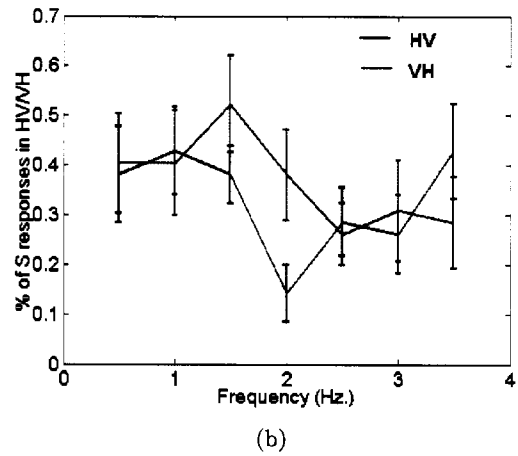
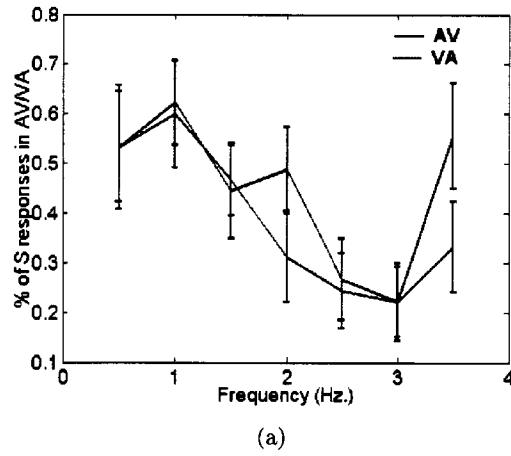


Figure 4.4: Shown is a measure of the symmetry of the S region as a function of frequency for the three different stimulus sequences. The black line indicates the number of S responses to the left of the $dT=0$ line and the blue line indicates the number of S responses to the right of the $dT=0$ line a) Auditory-visual stimuli. The black line is significantly higher for frequency = 2Hz and 3.5Hz indicating that the S region tilts towards the AV region. b) The black line indicates the tilt of the S region towards the HV region and the blue line indicates the tilt of S towards VH. c) The black line indicates the tilt of S towards HA and the blue line indicates a tilt towards AH.

time subsequent to the presentation of the visual stimulus. Beyond this window of convergence subjects perceive the sequential nature of the stimulus sequence. A parallel between the phenomenon of ‘temporal ventriloquism’ and the dynamics of excitable elements may be seen in the work of Jirsa and Kelso (2005). They demonstrated that the trajectories of two excitable elements coupled by a sigmoidal coupling term necessarily converge towards each other after a perturbation to each of the elements. A truncated form of the sigmoidal function, also known as the HKB coupling, results in convergent trajectories only if the perturbations to each element lie within a specific time window. Outside this range the two trajectories diverge from each other.

A second example of the parallels between the dynamics of an excitable system and multisensory perceptual phenomena is the flicker-flutter illusion. Shipley (1964) presented subjects with a rhythmic auditory and a rhythmic visual stimulus. Both stimulus sequences were initially set to the same frequency. On increasing the frequency of the auditory stimulus he found that the perception of the visual stimulus tended to follow that of the auditory stimulus though they differed in frequency. This occurred up to a certain frequency difference after which the perception of the stimulus sequence switched abruptly to asynchrony. Driving an excitable element with a rhythmic driver results in phase-locked oscillations within a certain frequency range beyond which phase locking is lost and multifrequency coordination and phase drift sets in (Kelso, 1995; Assisi et al., 2005a) akin to the driving of visual perception by a rhythmically presented auditory stimulus.

Here, we propose a model using coupled excitable elements and suggest a plausible coupling scheme between the individual sensory modalities and an interpretation of their interaction. The primary purpose of our model is to account for the following observations:

1. A temporal window of integration exists within which stimuli from two modalities are perceived as simultaneous.
2. This temporal window is asymmetric and depends on the modalities involved.
3. The width of the temporal window of integration is a function of the two different modalities involved.

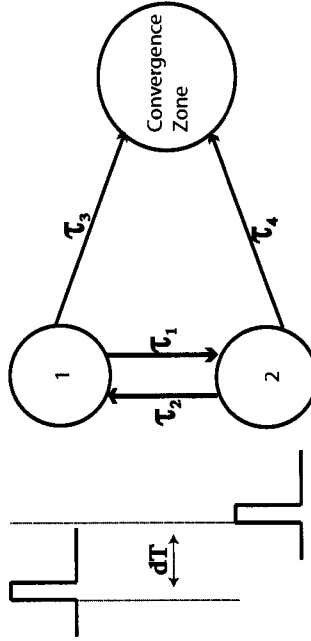


Figure 4.5: Shown is the architecture of the model implemented here. 1 and 2 are the two excitable elements coupled to each other. Temporal delays, τ_1 and τ_2 , are introduced into the coupling between the excitable elements. The delays τ_3 and τ_4 , between the excitable elements and the convergence zone is set to zero. The outputs from the two excitable elements are fed into a convergence zone that classifies the relative dynamics to a prescribed percept. See text for details.

4. There exists a region where the perception of simultaneity or asynchrony is disrupted giving rise to a new perceptual state, drift.

Our choice of model is motivated by the evidence from neurophysiology that suggests the existence of multisensory convergence zones whose role is to integrate information from different unisensory areas. Recent evidence has also suggested the existence of direct projections between the auditory and visual areas (Falchier et al., 2002; Rockland and Ojima, 2003). A plausible architecture that incorporates these two factors is shown in Figure 4.5.

In this two layered architecture, the individual sensory modalities represented by the elements 1 and 2 belong to the first layer that converges into a second layer. The ‘convergence

zone' classifies the dynamics of the first layer and gives rise to a particular percept. We do not employ a specific instantiation of the second layer. However, the classification scheme implemented by the second layer must possess the following properties,

1. If the response of one of the excitable elements converges toward the other then the ensuing perceptual response will be classified as simultaneous.
2. If the response of the two elements diverges the perceptual response will be classified as sequential or asynchronous.
3. If the response of the excitable elements to a rhythmic stimulus deviates from 1 : 1 phase-locked oscillations, the perceptual response will be classified as drift.

We now require an ideal candidate to describe the coupling between two excitable elements denoted by 1 and 2 in Figure 4.5. Various authors have used a simple step function to connect two neurons indicating that a postsynaptic neuron fires maximally or not at all depending on whether the potential of the presynaptic neuron is below or above threshold. For ensembles of neurons the mean activity may be similar to the step function described above, however deviations in the firing threshold may lead to a smoothed out step function resulting in a sigmoidal coupling function (Ermentrout, 1998; Wilson and Cowan, 1972). Additional effects of fatigue and attention would lead to a decline following the increase of the sigmoidal function (Freeman, 1992). This form of coupling obtained by truncating a sigmoidal nonlinearity to include only terms up to third order has been used extensively in the literature on sensorimotor coordination (Jirsa et al., 1998; Haken et al., 1985). The truncated sigmoidal coupling has been used to model pattern stability and phase transitions in bimanual coordination known as the Haken-Kelso-Bunz (HKB) model (Haken et al., 1985). The properties of convergence and divergence of the trajectories of excitable elements, given this particular form of coupling, is sufficiently general, in that it does not depend on the specific instantiation of the excitable elements (Jirsa and Kelso, 2005). In our analysis we choose the two-dimensional Fitzhugh-Nagumo neuron for each element in Figure 4.5 given by the following equations.

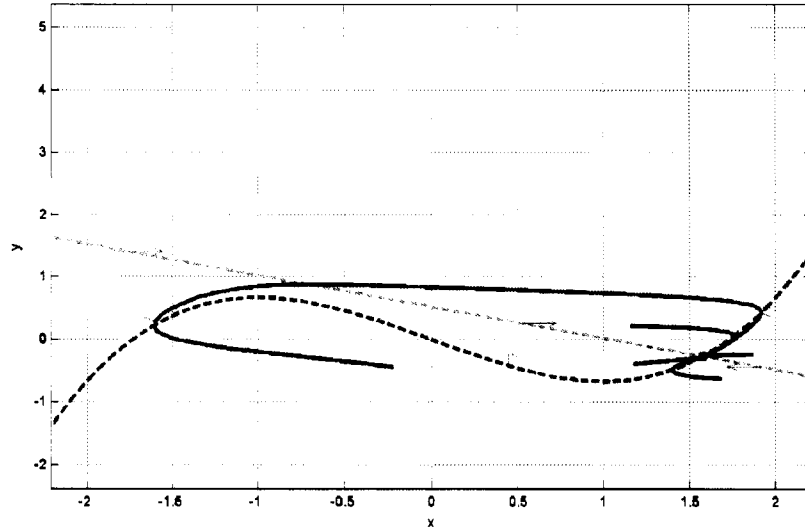


Figure 4.6: Phase space of the Fitzhugh-Nagumo equations. The dashed lines indicate the x and the y nullclines. The blue lines show different trajectories with different initial conditions. A small perturbation from the fixed point (intersection of the nullclines) causes the system to return after a small deviation. A large perturbation leads it along a long trajectory in phase space. Here the parameters are, $a = 0.7$, $b = 2$, $c = 3$. The parameter b determines the slope of the y nullcline given by the straight line in the figure.

$$\dot{x} = \left(x + y - \frac{x^3}{3}\right)\tau \quad (4.1)$$

$$\dot{y} = -\frac{1}{\tau}(x - a + by - I(t)) \quad (4.2)$$

The model reproduces two kinds of behavior. Following a brief supra-threshold perturbation of the y variable by an external input $I(t) > I_{thresh}$ the system moves along a large trajectory before settling to a fixed point. A sustained supra-threshold external input causes the system to oscillate at a frequency determined by the parameters a and b .

Figure 4.6 depicts the behavior of the model in the space of the x and y variables. The locus of points where the flow of the system is parallel to the y -axis is known as the x nullcline given by, $\dot{x} = 0$, and depicted by the cubic line (magenta) in Figure 4.6. The y nullcline, where the flow is parallel to the x -axis, is a linearly increasing function of x . The intersection of the two nullclines is a fixed point. On either side of the cubic nullcline the

flow of the system points in opposite directions. A small sub-threshold perturbation causes the system to settle back to the fixed point. However a supra-threshold perturbation in the x - direction causes the system to go along a long trajectory in phase space before settling to the fixed point. A notable feature of this system is the existence of two time scales, the fast time scale along the horizontal axis which is of the order of $1/\tau$ and a slow time scale of the variable y which is of the order τ . One can manipulate the time spent by the system on the slow manifold by adjusting the parameter b which affects the slope of the y nullcline. The closer the y nullcline is to being tangent to the peak of the cubic x nullcline, the slower the dynamics along the manifold.

We now examine the dynamics of two excitable elements coupled using a truncated sigmoidal coupling. The equations are given by,

$$\dot{x}_1 = (x_1 + y_1 - \frac{x_1^3}{3})\tau + (y_1(t) - y_2(t - \tau_2))(\alpha + \beta(x_1(t) - x_2(t - \tau_2)))^2 \quad (4.3)$$

$$\dot{y}_1 = -\frac{1}{\tau}(x_1 - a + by_1 - I_1(t)) \quad (4.4)$$

$$\dot{x}_2 = (x_2 + y_2 - \frac{x_2^3}{3})\tau + (y_2(t) - y_1(t - \tau_1))(\alpha + \beta(x_1(t - \tau_1) - x_2(t)))^2 \quad (4.5)$$

$$\dot{y}_2 = -\frac{1}{\tau}(x_2 - a + by_2 - I_2(t)) \quad (4.6)$$

The external input to the two excitable elements consists of a sequence of square wave pulses similar to those presented in the experiment. One sequence of rhythmic pulses $I_1(t)$ is applied to one of the elements while a second sequence of pulses $I_2(t)$ with the same frequency but a different phase is applied to the second element. Whether the responses of the two coupled elements are convergent or divergent is determined by comparison to a pair of uncoupled excitable elements. The difference ($t_{coupled}$) in the time between the response of element 1 to an external perturbation and the response of element 2 to a perturbation dT ms later was measured for the coupled and the uncoupled cases ($t_{uncoupled}$). If the difference $t_{coupled} - t_{uncoupled} > 0$ then the activity was said to be divergent and convergent otherwise. Figures 4.7 shows the time course of the two excitable elements which are classified as

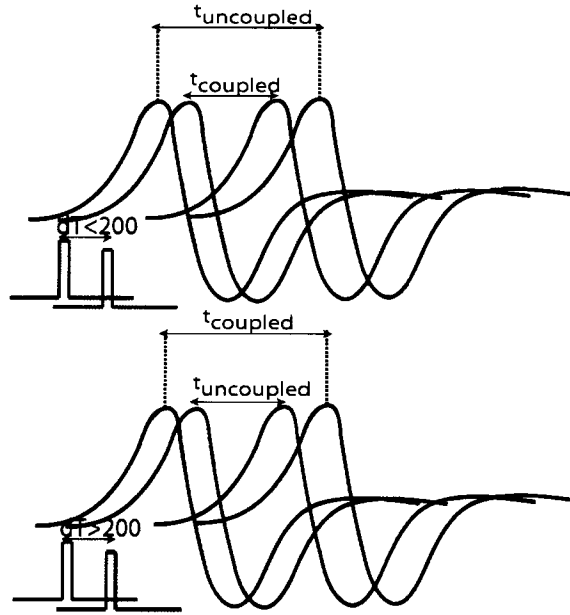


Figure 4.7: Shown are examples of convergent and divergent dynamics. Each trajectory is associated with an excitable element. The red trajectories indicate that the two elements are coupled while the black trajectories indicate that the two are uncoupled. a) An example of convergence. The trajectories of the two elements come closer to each other for the coupled case than for the uncoupled case. b) The trajectories for the coupled case diverge in comparison to that for the uncoupled case.

convergent or divergent. When the trajectories of the two elements approach each other the dynamics is classified as convergent and divergent otherwise.

Figure 4.8a-b illustrates the regions of convergence and divergence for the above pair of excitable elements. The colorbar indicates the value of $t_{coupled} - t_{uncoupled}$. The contour lines connect the points where $t_{coupled} - t_{uncoupled}$ are the same. Notice that the central region of convergence which is classified as the percept S is symmetric about the $dT = 0$ axis. In our model the width of the S region may be manipulated by changing the intrinsic time scale of each excitable element. For $b = 2$ the system spends a longer time on the slow manifold than when $b = 0$. We examined the width of the S region for two different values of b . The width of the S region associated with $b = 0$ is significantly smaller than the width of the S region for $b = 2$, that is, a faster intrinsic time scale ($b = 0$) is associated with a smaller region of simultaneity in comparison to slower intrinsic time scale ($b = 2$). A comparison between Figure 4.8a and Figure 4.8b demonstrates this fact. In our experiment we were able

to manipulate the symmetry and the widths of the perceptual partitioning by using different pairs of modalities. For the auditory-visual stimulus sequence the width of the region S is significantly larger than that for the haptic-auditory stimulus sequence (see Figures 4.2a-c and 4.3). We propose that one of the factors contributing to the width of the region S is the intrinsic time scales associated with the processing of stimuli from different modalities. The presence of a slow intrinsic time scale associated with the visual modality contributes to the larger width of the S region for auditory-visual stimulation in comparison to haptic-auditory stimulation. A second feature of the experiment captured by the model is the existence of asymmetries in the dT-ISI parameter space. In order to manipulate the symmetry of the system we introduce a time delay in the interaction between the two excitable elements. These time delays are introduced asymmetrically into the coupling with τ_1 , the time taken for a perturbation to travel from element one to element two, being smaller than τ_2 , the time taken for a perturbation to travel in the opposite direction. Figure 4.9 shows the behavior of the system simulated with the presence of an asymmetric delay term. A result of this asymmetry is that the region S tilts towards $dT < 0$. We observed this asymmetry in the perceptual partitioning of the parameter space associated with auditory-visual stimulation. In the model we choose to introduce delays only in the direct communication between the two excitable elements and not in its connections to the convergence zone, that is, $\tau_3 = \tau_4 = 0$. In our model there are no feedback connections from the convergence zone back to the excitable elements. Therefore any effects due to delays between the excitable elements and the convergence zone ($\tau_3, \tau_4 \neq 0$) will show up in terms of the difference $\tau_3 - \tau_4$. Also, since we have not specified a particular instantiation of the convergence zone and only required that it classify the inputs according to the rules mentioned earlier in this section, the effect of the delay $\tau_3 - \tau_4$ may be emulated by changing the value of the interonset interval, dT.

For frequencies higher than 0.06 Hz the 1 : 1 phase relationship between the stimulus and the response of each excitable element is lost. This region is classified as drift and corresponds to the drift region shown by the green contours in Figure 4.2. Our experimental results indicate that the region of drift is not uniformly spread out across all values of dT. There exists an asymmetric region of simultaneity for higher frequencies. Such segregated regions of synchrony suggests the presence of inherent nonlinearities in the system. This

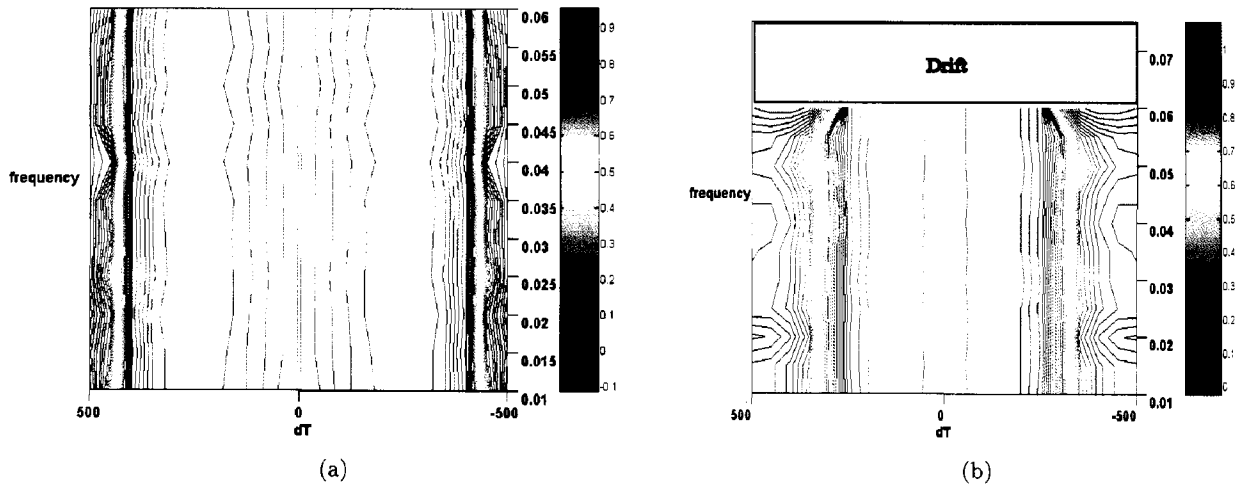


Figure 4.8: Contour plots showing the regions of convergence and divergence in the dT-ISI parameter space. The coupling is instantaneous. $\tau_1 = \tau_2 = 0$ a) Contour map using the parameter $b = 2$, indicating a slow intrinsic dynamics b) Contour map using the parameter $b = 0$, indicating a fast intrinsic dynamics.

The colormap indicates the degree of convergence or divergence increasing from blue to red. The values associated with the colors are given by $t_{coupled} - t_{uncoupled}$. Regions with blue contours indicate convergence, while red contours indicate divergence. The region where the 1:1 phase relationship between the stimulus and the response of the excitable element is lost is marked as ‘drift’.

particular feature of the experiment is not captured by our model. A possible scheme to explain the presence of this asymmetry is to introduce a higher order driving term. We demonstrated the role of higher order driving terms in chapter 3. A parametric driving term leads to the presence of isolated regions where coordinative states characterized by particular frequency ratios are stable (see chapter 3 for details).

4.4 Concluding remarks

In this paper we have presented a theoretical and experimental framework to discuss mechanisms of multisensory integration in the temporal domain. Key to our analysis has been the existence of different temporal scales associated with processing within and communication between different sensory modalities that lead to asymmetries in multisensory perception. We demonstrate that replacing one modality of stimulation with another does not alter

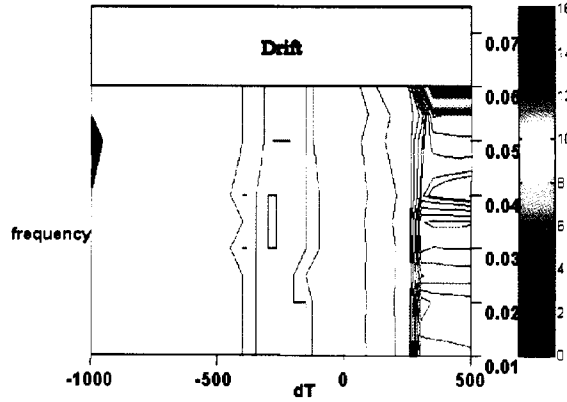


Figure 4.9: The contour map is similar to that shown in figure 4.8. The asymmetry is due to an asymmetric time delay. $\tau_1 = 0$ and $\tau_2 = 7ms$.

the qualitative features of the perceptual partitioning of the parameter space described above. This indicates a general mechanism of integration that is independent of the particular modalities involved. In this paper we provide a theoretical basis to discuss these issues and make an attempt to uncover the underlying modality independent mechanisms of integration. The model developed in section 4 uses a particular instantiation of an excitable system but is generalizable to other excitable systems as well (see Jirsa and Kelso (2005) for a detailed discussion of this issue). Interesting features of temporal multisensory integration reported in the literature include the phenomenon of ‘temporal ventriloquism’ (Morein-Zemir et al., 2003) where the perception of the visual stimulus is ‘ventriloquised’ towards a subsequently presented auditory stimulus. Using a different paradigm, Fendrich and Corballis (2001) demonstrate that this temporal cross-capture is an asymmetric phenomenon depending on the precedence of the stimuli. Our model captures the essential features of this temporal ventriloquism and its inherent asymmetry. A model of coupled excitable elements may also provide interesting clues into other multisensory illusions including the flicker-flutter illusion (Shipley, 1964), where the frequency of the visual percept follows that of the auditory percept though the presentation frequency of the auditory and visual stimuli may be different.

Chapter 5

Multisensory integration for timing engages different brain networks: An fMRI study

The possession of multiple sensory systems provides humans and other species with clear behavioral advantages. Not only can information from different modalities be used interchangeably, crossmodal integration substantially enhances our ability to detect, locate and discriminate external stimuli (for a review see Stein and Meredith (1993)). On the other hand, asynchronous, spatially disparate, and/or semantically incongruent inputs from different modalities are significantly less effective in eliciting responses than unimodal stimuli alone (Sekuler et al., 1997) and may degrade behavioral performance (Calvert, 2001). In the search for neural correlates of multisensory integration, neuroanatomical studies in primates have identified numerous areas where afferents from different senses converge. At the cortical level, such multisensory convergence zones include the superior temporal sulcus (STS), intraparietal sulcus (IPS), posterior insula and frontal regions including premotor, prefrontal and anterior cingulate (AC) (Jones and Powell, 1970; Seltzer and Pandya, 1989; Mesulam and Mufson, 1982). Subcortical anatomical convergence zones include the claustrum, the superior colliculus and the hippocampus (Mesulam and Mufson, 1982; Mufson and Mesulam, 1984; Pearson et al., 1982). Electrophysiological studies have shown that many

of these areas contain cells responsive to stimulation in more than one modality (Desimone and Cross, 1979; Vaadia et al., 1986; Duhamel et al., 1991). When two or more sensory cues from different modalities are congruent, i.e., appear in close temporal and spatial proximity, the firing rate of cells in multisensory convergence zones can increase multiplicatively, i.e. beyond that expected by summing impulses exhibited to each modality alone. By contrast, incongruent crossmodal cues can produce a profound response depression. Analogous evidence for auditory-visual integration sites using large-scale brain imaging has been found (Calvert, 2001; Bushara et al., 2001). Previous behavioral, anatomical and imaging studies suggest that congruent multisensory stimuli result in augmented behavior and enhanced neural responses in multisensory convergence zones. Incongruent stimuli, on the other hand, result in degraded behavioral performance and neural response depression in multisensory convergence zones. Here we wish not only to explore the neural correlates of the percepts of congruent and incongruent stimuli, but also to establish a neutral percept, which represents the failure to integrate multisensory stimuli. To this end we develop a rhythmic multisensory paradigm, in which temporally congruent stimuli cause a percept of synchrony and temporally incongruent stimuli cause a percept of asynchrony. A beauty of the rhythmic paradigm is that a third possible perceptual state is possible, the non-phase locked state, also known as phase wrapping or drift (Kelso, 1995). Here our behavioral studies show that subjects perceive the multisensory stimulus sequences as two co-existing entities with no fixed phase relation (see chapter 4 for a detailed exposition). This paradigm thus affords not only a comparison of dichotomous conditions corresponding to congruent and incongruent stimuli, but also allows us to ask the more subtle question of which brain networks underlie the formation and dissolution of multimodal percepts as stimulus parameters change. In our experiment, the timing parameter space consisted of intra-stimulus intervals varied from -200 ms (light before sound) to +200 ms and stimulation rates (or the inverse of interstimulus interval) from 0 to 3.5 Hertz, as shown in Fig. 5.1. 16 subjects participated in the behavioral experiment (see Appendix B for details). They were instructed to perceive bimodal stimuli as synchronized events, maintaining a constant level of attention toward either modality. Their responses indicated that the perception of auditory-visual stimuli changed qualitatively according to the values of timing parameters. As shown in Fig 5.1, there were four

distinct perceptions: perception of (i) synchrony (S), (ii) auditory leading visual stimuli (AV), (iii) visual leading auditory stimuli (VA), and (iv) drifting order of stimuli, in which subjects could report no clear percept (D). Notice in Fig. 5.1 the perceptual region of synchrony is asymmetrical around the intra-stimulus interval, extended more toward auditory leading visual stimuli ($p < 0.02$, shown in Fig 5.1). Below the stimulation rates of 2.5 Hz, the perceptions of synchrony and asynchrony persist, whereas above 2.0 Hz a region of no fixed percept predominates. The existence of these distinct perceptual states in the physical timing parameter space indicates the definite role of inherent delays involved in multisensory processing.

To find the neural correlates of these four perceptual states, we scanned 13 subjects who had demonstrated consistent perceptions in the selected regions (dark grids in Fig. 5.1b) of the timing parameter space. The imaging experiment consisted of three functional runs of sensory (auditory, visual and auditory-visual) stimulation and rest conditions in an on-off block design. Auditory and visual stimuli were presented in the scanner through a pair of headphones and goggles respectively. The data were preprocessed and analyzed using SPM2. (see Appendix A for a detailed description of the methods.) The data from all thirteen subjects were combined in a random effects analysis to identify brain regions of activations common in all bimodal conditions and in each condition. The region of interest (ROI) analysis was performed by extracting time courses from the regions of activations to determine the fMRI signal change during on-off blocks.

Nine of the 13 subjects were consistently able to establish bistable percepts during the bimodal conditions. Results showed bilateral activations at $p < 0.001$ in the inferior frontal gyrus (I), superior temporal gyrus (II), middle occipital gyrus (III) and inferior parietal lobule (IV) (a sagittal slice at $x = -47.75$ shown in Fig. 5.2. Interestingly, a negative contrast between bimodal conditions relative to rest was revealed in the posterior midbrain in the region of the superior colliculus (V), as shown in the coronal slice at $y = -30$ in Fig. 5.2. These results indicate that the network for multisensory processing of auditory-visual events is composed of a collection of areas not all of which were seen to be significantly active in separate individual conditions.

Among significantly activated areas for individual bimodal conditions is the left inferior

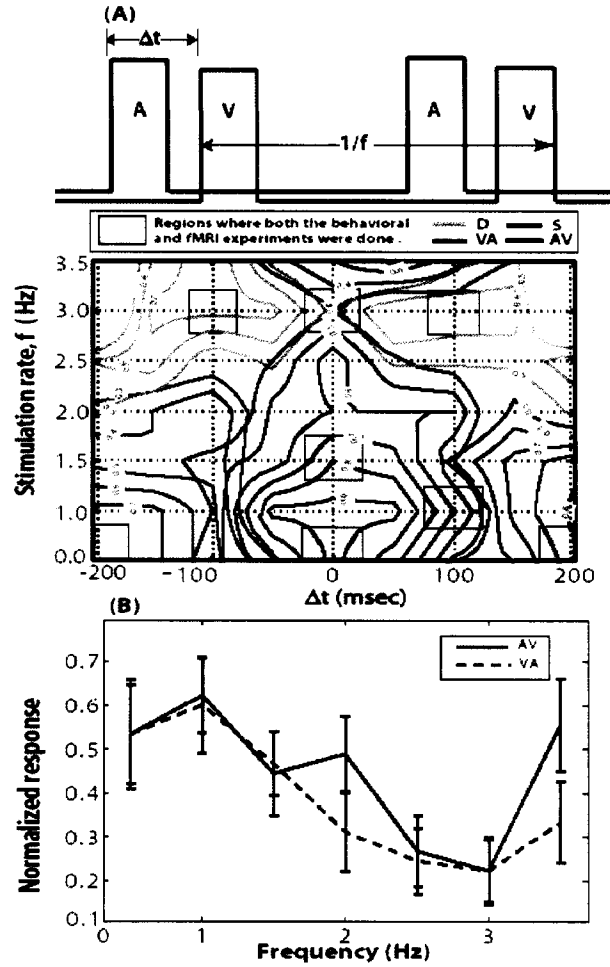


Figure 5.1: (A). Experimental design. A visual (V) stimulus is presented after an auditory stimulus (A) a time interval Δt (intra-stimulus interval) and repeated at a rate f (stimulation rate). Behavioral response in the space of Δt and f : the negative values of Δt imply that the visual stimulus (V) precedes the auditory stimulus (A) and the contour levels represent the normalized response (number of responses divided by the total possible responses). Depending on these timing parameters, the participants reported the perceptions of AV (sound before light), VA, S (synchronous) and D (drifting order), thereby partitioning the space into four distinct regions. Notice the asymmetry of S-region, extended more toward the region of AV.

(B). Mean normalized response versus stimulation rate for the perception of synchrony in the regions of VA and AV. The response in both regions first decreases and then increases with stimulation rate. The errorbars are the standard error mean. There is a significant asymmetry, extended more toward the region of AV ($p < 0.02$).

Note that the figure shown above is the same as Fig 4.4a

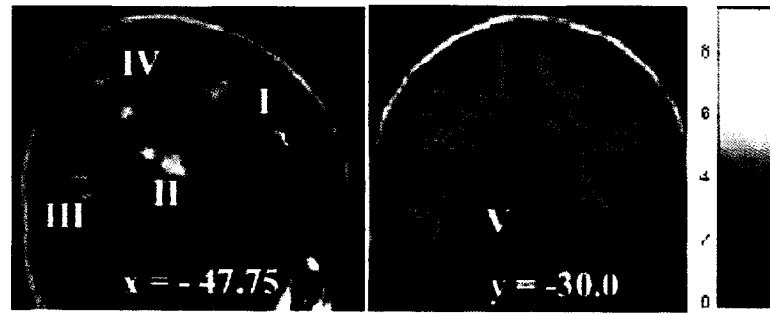


Figure 5.2: Activations related to crossmodal processing ($p < 0.001$). Areas: inferior frontal (I), superior temporal gyrus (II), middle occipital gyrus (III), inferior parietal lobule (IV), and posterior midbrain (V). The activation in the posterior midbrain (V) in the region of superior colliculus was obtained by a negative contrast of task versus rest. The color intensity represents t-statistics and the activations are overlaid on the Montreal Neurological Institute (MNI) structural template brain in the neurological orientation for display of the t-maps.

parietal lobule for which the stimuli were perceived as asynchronous (± 200 ms, 0.5 Hz), (-200 ms, 0.5 Hz) and (± 100 ms, 1.0 Hz) (shown in Fig. 5.3). The associated mean time courses from the left inferior parietal lobule also showed significantly greater signal change (0.3 %, $p < 0.01$). Activation of the inferior parietal lobule at (± 100 ms, 1.0 Hz) corresponded to the percept of audition preceding vision. However, in the perceptual drift regime at (± 100 ms, 3.0 Hz), there was no significant activation in the inferior parietal cortex.

From some previous studies of non-rhythmic tasks (Bushara et al., 2001; Assmus et al., 2003), the inferior parietal cortex has been shown to be involved in the detection of asynchrony and in integrating spatial and temporal information. Our results indicate that a network composed of frontal, auditory, visual and inferior parietal areas is crucial for the formation of this percept. On the other hand, the disengagement of the inferior parietal areas from the activated network was related to the perceptual drift. Nine subjects were able to establish the percept of synchrony between auditory and visual stimuli for the timing parameters (± 100 ms, 1.0 Hz). A negative contrast of task versus rest ¹ revealed activation in the posterior midbrain in the region of superior colliculus shown in Fig. 5.3. The supe-

¹The brain activations associated with a negative contrast indicate the involvement of the area with a higher baseline of activity. The timeseries from fMRI signal change extracted from the cuneus and posterior midbrain showed a higher signal change for off-condition than for on-condition. Especially at the time of switching from task to no task the most of the fMRI signals from these locations started to increase and then decreased for about 9 seconds as in an event. This switching from task to no task involved a change in visual scene from a red flash to a fixation cross.

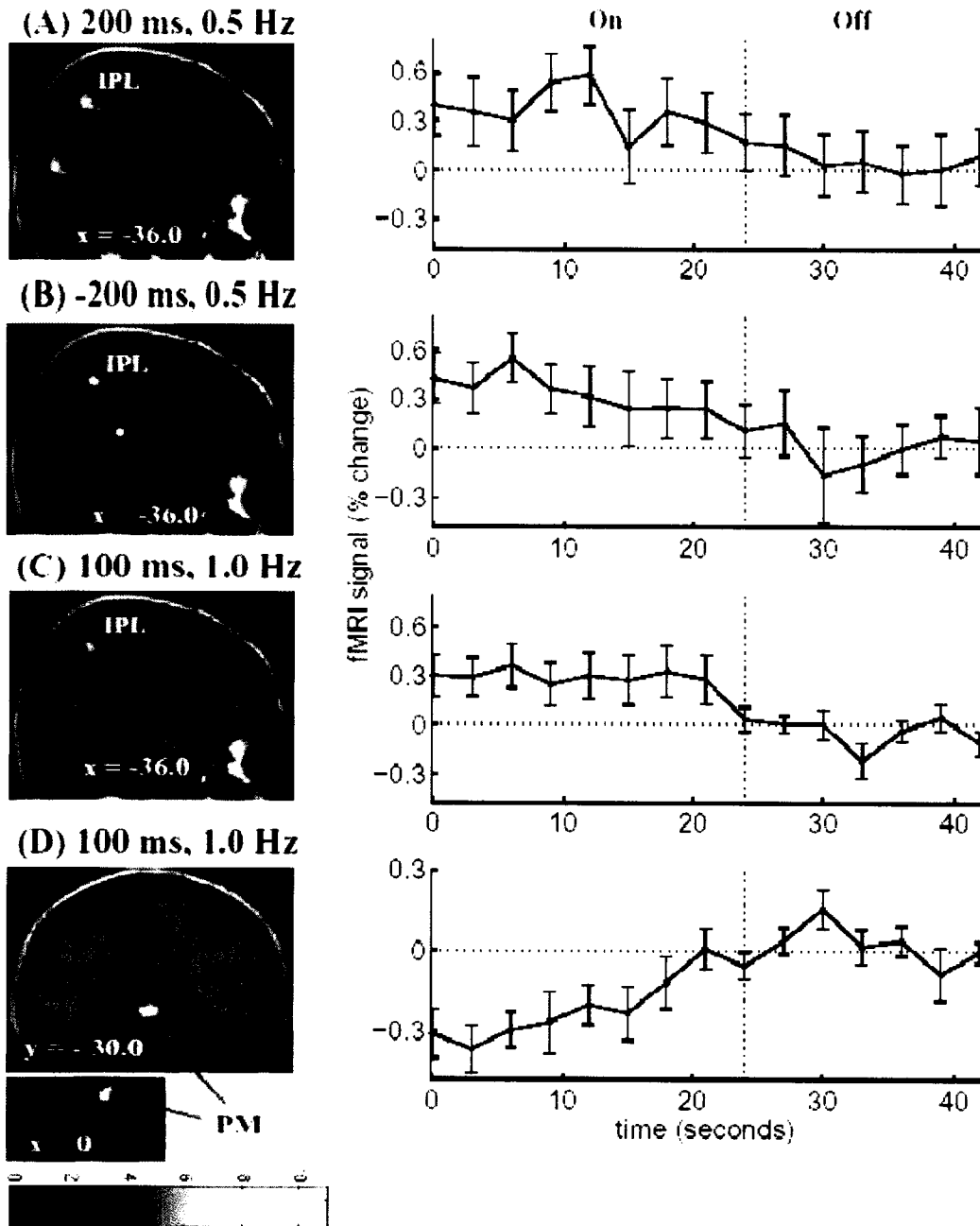


Figure 5.3: (A-C). Activations related with the perception of asynchrony ($p < 0.001$) and the associated mean BOLD responses for the following conditions: (A) ($\Delta t, f$) = (200 ms, 0.5 Hz), (B) (-200 ms, 0.5 Hz) (C) (100 ms, 1.0 Hz). The inferior parietal lobule (IPL) was active as an area of the sub-networks for this percept. The overall signal change between the task and the rest is about 0.3 % ($p < 0.01$). The error bars represent the standard error mean. All subjects were able to establish the perception of asynchrony except for the condition (100 ms, 1.0 Hz), where only nine subjects were able to do that in the scanner. On the other hand, the inferior parietal cortex was not active for (Δt 100 ms, 3.0 Hz), where there was no fixed percept. (D) Activation associated with the perception of synchrony ($p < 0.001$) and the associated mean BOLD response. The negative contrast (18) of the task versus rest activated the posterior midbrain (PM) in the region of superior colliculus for the condition (100ms, 1.0 Hz). The subjects were instructed to get the perception of synchrony and nine subjects could establish the percept. Here, the time series increases continuously during the on-condition unto 6 seconds after the stimuli were turned off and then decreases

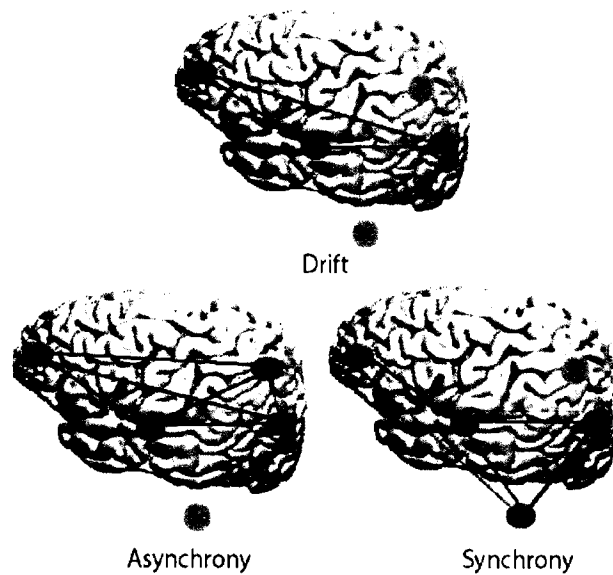


Figure 5.4: Illustrative diagram of the sub-network of active cortical and subcortical brain areas (indicated in red) for crossmodal processing in the perception of asynchrony, synchrony and when there was no clear percept (drift). Based on the brain activations (19) under bimodal conditions, the perception of asynchrony is associated with the sub-network consisting of frontal (F), auditory (A), visual (V), and parietal (P) areas, the perception of synchrony with the sub-network consisting of frontal (F), auditory (A), superior colliculus (Sc) and visual (V) areas. When there was no clear percept formation, active sub-network consisted of frontal (F), auditory (A) and visual (V) areas. (Here, the cortical surface of the MNI brain and MNI single brain template have been used with markers (red and green) to indicate approximate locations for illustrative purposes only).

rior colliculus is best known to possess multisensory neurons that likely transform multiple sensory signals into purposeful action Stein and Meredith (1993). Once again, a network of areas is seen to support the perception of auditory-visual synchrony. Only this time the network includes the superior colliculus and not the inferior parietal lobule. Figure 5.4 summarizes our results using a schematic diagram to illustrate the networks activated during the perceptions of asynchrony, synchrony and drift and how the components of the network change and reorganize with perceptual change.

In conclusion, our findings show that different distributed networks are activated when perceptually synchronous and asynchronous stimuli are processed. In particular, the superior colliculus is associated with the perception of synchrony and the inferior parietal cortex with the perception of asynchrony of auditory-visual signals. However, when there is no clear

percept, these network components are disengaged and only a residual network comprised of sensory and frontal areas remains active. These results demonstrate that the processes of perceptual integration and segregation engage and disengage different brain subnetworks, but leave only a rudimentary network in place if no clear percept is formed.

Chapter 6

Spatial and temporal modulations of multisensory source activity: An EEG study

6.1 Introduction

Our understanding of cortical multisensory integration has been significantly influenced by the hierarchical model of sensory processing. Environmental influences from a given modality undergo extensive processing with the simpler features being processed at earlier stages followed by progressively more complex features. In a recent review Mesulam (1998) describes a subdivision of the cortex into five major subtypes: primary sensory-motor, unimodal association cortex, heteromodal association, paralimbic and limbic regions. This functional architecture derives primarily from anatomical, physiological, and behavioral experiments on macaque monkeys and suggests a sequential, convergent organization with progressively more complex functions as one ascends along the hierarchy (Mesulam, 1998; Falchier et al., 2002; Seltzer and Pandya, 1989). Based on this architecture environmental stimuli are assumed to undergo extensive unimodal processing before they converge into multisensory convergence zones. Neuroanatomical studies in primates have identified numerous areas where afferents from different senses converge. At the cortical level, these multisensory

convergence zones include the superior temporal sulcus (STS), intraparietal sulcus (IPS), posterior insula and frontal regions including premotor, prefrontal and anterior cingulate (AC) (Jones and Powell, 1970; Seltzer and Pandya, 1989; Mesulam and Mufson, 1982). Subcortical anatomical convergence zones include the claustrum, the superior colliculus and the hippocampus (Mesulam and Mufson, 1982; Pearson et al., 1982; Mufson and Mesulam, 1984). Electrophysiological studies have shown that many of these areas contain cells responsive to stimulation in more than one modality (Desimone and Cross, 1979; Vaadia et al., 1986). In particular, the cells in the superior colliculus have been investigated in detail with respect to crossmodal integration (Stein and Meredith, 1993). Evidence for auditory-visual integration sites using large-scale brain imaging has been found in the superior temporal sulcus using speech and fMRI, the superior colliculus (Calvert et al., 2001) using non-speech stimuli and fMRI and intraparietal sulcus (IPS) (Bushara et al., 2001) using non-speech stimuli and PET. These multisensory convergence zones have been identified primarily in the heteromodal association areas, which in Mesulam's account are three synapses away from the primary sensory areas. The geographical location of these convergence zones leads us to believe that the influences of multisensory integration must be seen rather late in the sequence of sensory processing. This led Calvert et al. (1999) to propose that the modulation of the activity in the primary sensory cortices coincident with multisensory stimulation must be due to feedback processes rather than feedforward mechanisms. Recent evidence from the electrophysiological and the anatomical literature has challenged this view. Giard and Peronnet (1999) have demonstrated in an ERP study that multisensory influences occur early (~ 40 ms) in sensory processing. They base their conclusions on the timing and topography of the C1 potential, presumed to be earliest cortical visually evoked potential. The findings of their study were corroborated by Molholm et al. (2002) using a higher density sensor coverage and simpler stimuli.

The prevalent approach in determining multisensory influences using EEG has been to employ what is known as the summative model. The model states that neural activities induced by the multisensory stimuli (AV) may be expressed as a sum of the neural activities due to unisensory stimulation ($A + V$) in addition to the neuronal activities induced uniquely by multimodal stimulation. This model has been applied within a temporal window

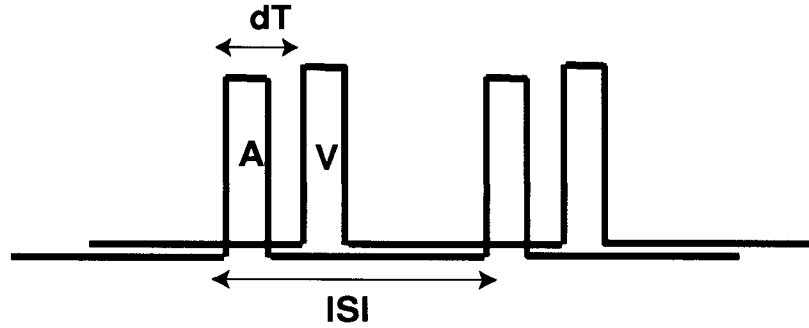


Figure 6.1: An example of the stimulus sequence and the parameters of the experiment. Similar stimulus sequences were used for Haptic-Visual and Haptic-Auditory stimulation.

that lasts 200ms post stimulus under the assumption that unimodal ERP's reflect sensory analysis in the modality specific areas. The residual activity, that is, the activity unique to multimodal stimulation given by, $R_{AV} = AV - (A + V)$, may be due to the following two factors: i) Temporal modulation of the unisensory components ii) The addition of a new source of activity that is different from the generators of the unimodal responses.

Overwhelming evidence suggests that multisensory influences begin to act at early stages of sensory processing and influence what has been traditionally considered unimodal areas (Rockland and Ojima, 2003; Rockland, 2004; Falchier et al., 2002). In order to understand the nature of this influence it is important to distinguish between the two factors mentioned above. In this chapter we derive a technique based on spatiotemporal mode decompositions that allows us to make this distinction. Subsequently, we use this technique to examine a novel experimental paradigm designed to examine the temporal features of multisensory integration parametrically.

6.2 Experimental design

We developed a generalized synchronization paradigm that is designed to test temporal features of multisensory integration parametrically. The stimulus sequence consisted of a pair of rhythmically presented auditory-visual stimuli. The parameters varied during the experiment were the interval between the onsets of the two stimuli of a pair, dT , and the inter-stimulus interval (ISI).

This design leads to a two dimensional parameter space. Each trial of the experiment samples one point in the dT-ISI parameter space. Subjects were required to report their perception of the stimulus sequence according to a forced choice paradigm. The reported perception of the subject for every sampled point in this space was used to partition the parameter space. In order to get a broad overview of the perceptual partitioning of the parameter space and its asymmetries, we first conducted behavioral experiments the results of which guided our subsequent EEG experiments. Subjects were instructed to try to perceive the multisensory stimuli as synchronous events. We indicated that this would not be possible across all experimental conditions. At the end of each trial subjects were required to report their perception of the stimulus using one of four options.

1. The auditory stimulus precedes the visual stimulus. (AV)
2. The visual stimulus precedes the auditory stimulus. (VA)
3. The two stimuli appear simultaneous. (S)
4. Can't tell. (D)

Particular mention must be made of option 4. This option corresponded to an auditory-visual illusion where the perceived temporal order of the stimulus changed across the length of the trial. An overview of the results is shown in Figure 6.2.

The inter-onset interval (dT) is plotted along the x-axis and the different frequencies of stimulation ($1/ISI$) are plotted along the y-axis. In Figure 6.2, negative values of dT indicate that the auditory stimulus preceded the visual stimulus. The maps in the Figure show the probability that a given subject reported a particular percept for a given dT-ISI point. The probability is given by the ratio of the number of subjects who reported a particular percept at a point in the two dimensional parameter space to the total number of subjects. The different perceptual states are shown in different colors according to the legend in the Figures. The contour lines for each percept connect points of equal probability in the dT-ISI space. The contour lines are truncated for probabilities below 0.3. The colored squares superimposed over the contours indicate the points in the parameter space where the EEG experiment was conducted. Each of these conditions correspond to different perceptions

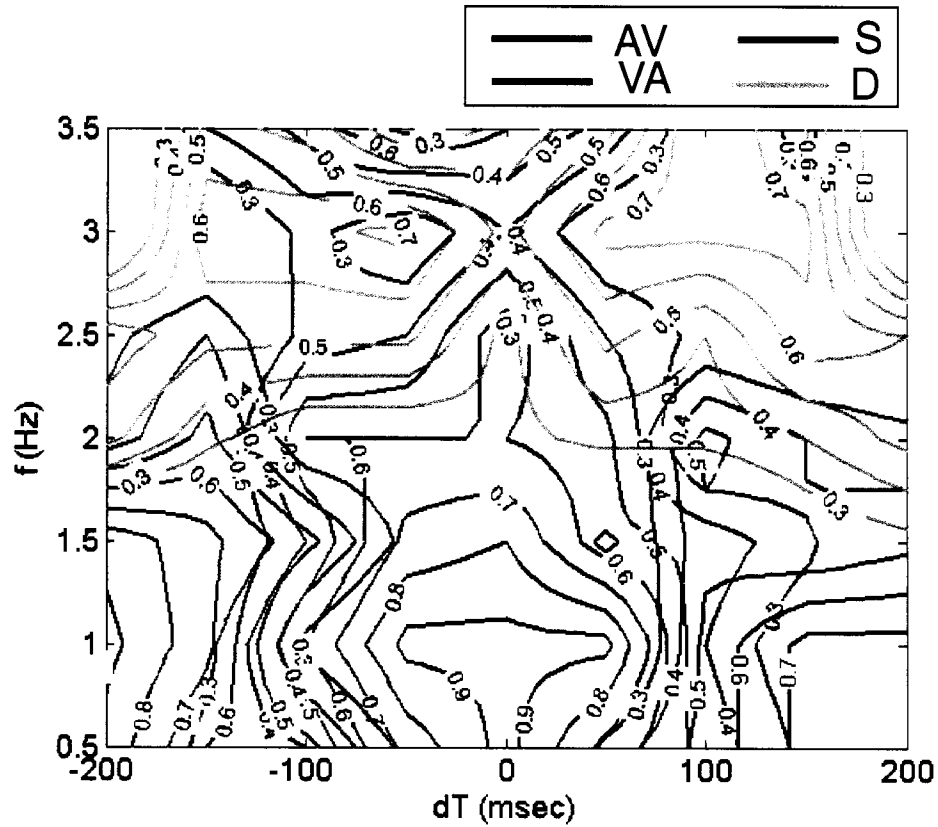


Figure 6.2: Overview of the behavioral experiment. The contours depict the probability that a subject reports a particular percept for a given point in dT-ISI space. The boxes represent the stimulus conditions that were presented to subjects in the EEG experiment. The red box indicates a condition where we compared the subjects response to two different sets of instructions without changing the physical stimulus (See text for details).

of the stimulus. The red square indicates an experimental condition that was unique to the EEG experiment. We noticed that the perception of the stimulus in this region of the parameter space depended on the instructions given to subjects. They could perceive it as either simultaneous (S) or sequential (AV). We conducted two sets of trials on subjects, one when they were instructed to try to synchronize the stimuli and the other when they were asked to perceive it as AV.

In Chapter 4, we obtained the neural correlates associated with the percepts of synchrony (S), asynchrony (AV and VA) and drift (D). Though we were able to determine a set of sources that are associated with these percepts, the low temporal resolution of fMRI constrains our interpretation of the temporal modulation of these sources. In subsequent sections we will be interested in the temporal features of multisensory processing, in particular, 1) How are the temporal features of a unisensory subsystem (say visual) modulated due to a simultaneously presented stimulus belonging to a different modality (say auditory) ? 2) What is the time course of the activity of uniquely multisensory sources, that is, sources that are activated only in response to multisensory stimulation but not to stimulation in a single modality?

Before we get into the technical details of experimental implementation and the specifics of our data analysis methods we discuss our motivation to address the above questions.

6.2.1 Temporal modulations of the unisensory sources due to multisensory stimulation.

A broad overview of the results of the behavioral experiment is shown in Figure 6.2. The same behavioral experiment was also conducted using other stimulus pairs, that is, haptic-auditory and haptic-visual (see chapter 3 for details). We demonstrated that the main features of the perceptual partitioning were unchanged regardless of the specific stimulus pair, suggesting that the features are not modality specific, but rather a signature of the couplings involved. Recent anatomical evidence has suggested that some of these couplings occur at early stages of sensory processing (Rockland and Ojima, 2001, 2003; Falchier et al., 2002). It has been demonstrated that the earliest effect of a visual stimulus manifests itself in the primary visual cortex at a latency of approximately 46ms (Foxe and Simpson,

2002; Foxe and Schroeder, 2005). Any modifications of this response when an auditory stimulus is presented simultaneously, must then be attributed to the interaction between the auditory and the visual subsystems. Whether this interaction is due to direct anatomical connections between the unisensory areas, or some other pathway is still an open question. It is important, however, to examine whether there is a modification of the response of the visual areas due to the simultaneous presentation of an auditory stimulus. In particular we will examine the condition $dT=0$, $f=0.5$ Hz for signs of these early interactions. Modifications in the activity of the primary visual areas will manifest itself as scalp potentials with an occipital peak and latencies ranging from 30ms to 60ms.

6.2.2 Time course of uniquely multimodal sources.

Our fMRI experiment (See Chapter 4) demonstrated that the sources activated in response to multisensory stimulation consisted of primary sensory, parietal and frontal areas in addition to the posterior midbrain. Different percepts were associated with different sub-networks. The perception of asynchrony (AV and VA) depended on a sub-network that included the inferior parietal cortex. Synchronous percepts (S), on the other hand, resulted in a disengagement of the inferior parietal cortex and the recruitment of the superior colliculus. For the percept of drift (D), only prefrontal and sensory areas were active. Of particular interest to our EEG experiment, is the condition marked by the red box in Figure 6.2. Here, the same physical stimulus pair generates two different percepts, that of synchrony and asynchrony, depending on the instructions given to the subject. In our EEG experiment this condition provides an ideal comparison between the percepts of synchrony and asynchrony. The fMRI study suggests that one of the differences between these two percepts is the preferential activation of the inferior parietal lobule for the percept of synchrony. We expect to see a signature of this activation in the topography of potentials measured on the scalp.

The percept of drift, leads to the deactivation of the posterior midbrain regions and the inferior parietal areas. We expect that this deactivation must lead to an overall decrease in the power associated with the activity of uniquely multimodal sources.

In section 6.4 we will develop a methodology that allows us to separate the contributions

to multisensory processing due to the temporal modulations of unisensory sources (section 6.2.1) and the activity of uniquely multisensory sources (section 6.2.2). We compare our approach to the summative model that has been used previously in the literature (Giard and Peronnet, 1999; Molholm et al., 2002).

6.3 Methods

6.3.1 Subjects

Four right handed male subjects participated in the experiment. Subjects had normal or corrected to normal vision and reported no hearing deficiencies. Informed consent was collected from each subject prior to the experiment, and the study was approved by the Florida Atlantic University Institutional Review Board.

6.3.2 Procedure

Auditory and visual stimuli were used during the experiment. A total of nine experimental conditions were performed. Each condition was characterized by a different value of the inter-onset interval between the two stimuli of a pair (dT) and the presentation rate ($f = 1/\text{ISI}$). The experiment was divided into 8 multisensory trials, and 6 unisensory control conditions each lasting a minute in duration. The number of trials required for a particular experimental condition was determined by the requirement that each experimental condition must have at least 250 stimulus pairs. The specific stimulus sequence and instructions given to subjects is detailed in the previous section. In addition to the trials shown in Figure 6.2 the unisensory control conditions were also conducted using auditory-only and visual-only stimulation at frequencies of 0.5, 1.5 and 3.0 Hz.

6.3.3 Data Acquisition

EEG activity was recorded using an 84-channel MANSCAN system (Sam Technology Inc., San Francisco, CA.). The electrodes were placed on the scalp using an Electrocap recording cap with the electrodes placed according to an extension 10/20 International system. The Manscan system uses a linked mastoid reference for recording. The signal was digitized at

256Hz. Immediately after the experiment the 3D location of the electrodes was determined using a Polhemus Isotrack II (Colchester, VT) digitizer.

6.3.4 Preliminary data analysis

The data were inspected visually for eye blink, muscle and movement artifacts which were marked in the data. The initial data inspection and cleaning was done using the MANSCAN software. All further data analysis was done using MATLAB. Each trial was first divided into epochs. Epochs containing artifacts were eliminated from all further analysis. The length of an epoch depended on the frequency of stimulus presentation. (2000ms for 0.5Hz, 667 ms for 1.5Hz and 330ms for 3Hz stimulation rate.) Each epoch was centered at the onset of the first stimulus of a multisensory stimulus pair. For the unisensory control conditions the epoch was centered at the onset of the auditory or the visual stimulus. The responses were averaged across epochs to obtain the event related potential. All further data analysis was done on the averaged signals.

6.4 Mathematical formulation of the method of analysis.

In this section we will evolve a mathematical formulation that allows us to differentiate between the contributions to multisensory processing due to the temporal modulation of the unisensory sources and the activation of sources unique to multisensory stimulation. In our experiment we measure potentials on the scalp that are a signature of the activity of sources in the brain. We will develop our methodology by first relating the measured scalp potentials to the sources in the brain (section 6.4.1) and apply this understanding to the specific multisensory paradigm of interest (section 6.4.2). Finally, we will use the adjoint basis decomposition (section 6.4.3) to extract the time series associated with the unisensory and the uniquely multisensory sources.

6.4.1 Relationship between the brain sources and the measured potentials.

Consider a set of m sources in the brain that are responsive to the presentation of an auditory stimulus. Assume that each of these sources may be approximated by a current dipole located at a particular position in space, $\mathbf{X}=(x,y,z)$, and oriented in the direction $\Theta = (\theta, \phi)$. The potentials measured on the scalp using EEG are superpositions of the activity of these sources after spatial filtering by the skull and the layers of tissue that lie between the sources and the recording electrodes on the scalp. Assume that there are n recording electrodes. The potentials measured on the scalp may, in general, be written as follows

$$\mathbf{A}(t) = \mathbf{H}_\mathbf{A}(\mathbf{X}, \Theta) \mathbf{a}(t) \quad (6.1)$$

where $\mathbf{a}(t) = [a_1(t) \cdots a_m(t)]^T$ are the underlying source activities as a function of time and $\mathbf{A}(t) = [A_1(t) \cdots A_n(t)]$, are the potentials measured on the scalp at time t . The matrix $\mathbf{H}_\mathbf{A}(\mathbf{X}, \Theta)$ is an $n \times m$ time independent transformation matrix that, in addition to the location and orientation of the sources, contains information about physical properties of the skull and tissue between the sources and the sensors. The collective activity of the sources may be decomposed into time dependent amplitudes and time independent modes as follows

$$\begin{bmatrix} a_1(t) \\ \vdots \\ a_m(t) \end{bmatrix} = \eta_a^1(t) \pi_a^1 + \cdots + \eta_a^l(t) \pi_a^l \quad (6.2)$$

In general, $l \leq m$. A total of m modes is always sufficient to completely capture the dynamics of all the sources. However, consider a limiting case, where the activity of two of the m sources is perfectly locked in time. Then the system may be completely defined by a set of $m - 1$ modes and their corresponding time dependent amplitudes. In general, we may infer that each mode is representative of the simultaneous activity of a set of sources. Importantly, this does not preclude the possibility that a given set of sources may contribute

to multiple modes. The matrix, $\mathbf{H}_A(\mathbf{X}, \Theta)$, relates the activity of the sources in the brain to the potentials measured on the scalp. It acts as a spatial filter that blurs the distinction between nearby sources therefore decreasing the dimensionality of the source space. The operation of the spatial filter on the sources may thus be written as,

$$\mathbf{H}_A(\mathbf{X}, \Theta) \{ \eta(t) \pi_a^1 + \dots + \eta_a^l(t) \pi_a^l \} = \tilde{\xi}_A^1(t) \tilde{\phi}_A^1 + \dots + \tilde{\xi}_A^k(t) \tilde{\phi}_A^k \quad (6.3)$$

The decomposition given by equation 6.3 represents the spatiotemporal behavior of the potentials measured on the scalp and may be captured by a set of k modes. As mentioned earlier, the matrix $\mathbf{H}(\mathbf{X}, \Theta)$ acts as a spatial filter. Therefore, a smaller number of modes, k , will be required to describe the spatiotemporal behavior of the potentials measured on the scalp than the number of modes required to describe the source activity, l , that is, $k \leq l \leq m$. The modes, $\{\tilde{\phi}_A^i : i = 1, \dots, k\}$ represent the activity of a sets of sources. Any non-singular linear transformation of the above modes results in a rotation and scaling of the modes, thus defining a new coordinate system. The interpretation that each of the rotated and scaled modes, which we designate $\{\phi_A^i : i = 1, \dots, k\}$, results from the activity of a particular set of sources still stands. In general, we may obtain a spatial mode decomposition of the potentials measured on the scalp using a number of different methods. These modes, $\{\phi_A^i : i = 1, \dots, k\}$, may be seen as a linear transformation of $\{\tilde{\phi}_A^i : i = 1, \dots, k\}$.

6.4.2 Differentiating between the spatial and temporal features of multi-sensory processing.

The scalp potentials measured in response to auditory stimulation may be expressed as,

$$\begin{bmatrix} A_1(t) \\ \vdots \\ A_n(t) \end{bmatrix} = \xi_A^1(t) \phi_A^1 + \dots + \xi_A^k(t) \phi_A^k \quad (6.4)$$

Similarly, the scalp potentials measured in response to a visual stimulus may be expressed as follows,

$$\begin{bmatrix} V_1(t) \\ \vdots \\ V_n(t) \end{bmatrix} = \xi_V^1(t)\phi_V^1 + \dots + \xi_V^{k'}(t)\phi_V^{k'} \quad (6.5)$$

The summative model assumes that the scalp potentials obtained in response to auditory-visual stimulation may be expressed as a sum of the auditory-only and visual-only responses. In terms of equations 6.4 and 6.5 the summative model may be written as follows,

$$\begin{bmatrix} AV_1(t) \\ \vdots \\ AV_n(t) \end{bmatrix} = \sum_{i=1}^k \xi_A^i(t)\phi_A^i + \sum_{j=1}^{k'} \xi_V^j(t)\phi_V^j + R_{AV}(t) \quad (6.6)$$

where the residual term $R_{AV}(t)$ is that part of the data not explained by the summed unisensory contributions. We now wish to express the scalp potentials due to auditory-visual stimulation in terms of auditory only and visual only stimulation by projecting the multisensory data onto the unisensory modes. This may be represented as follows,

$$\begin{bmatrix} AV_1(t) \\ \vdots \\ AV_n(t) \end{bmatrix} = \sum_{i=1}^k \xi_{AV_A}^i(t)\phi_A^i + \sum_{j=1}^{k'} \xi_{AV_V}^j(t)\phi_V^j + R_{AV_{spatial}}(t) \quad (6.7)$$

The modes ϕ_A^i and ϕ_V^j , as mentioned earlier, are associated with the set of sources that are simultaneously activated in response to auditory-only and visual-only stimulation respectively. Henceforth, we will refer to the modes $\{\phi_A^i : i = 1, \dots, k\}$ as the auditory subsystem and the modes $\{\phi_V^j : j = 1, \dots, k'\}$ as the visual subsystem. If the residual $|R_{AV_{spatial}}(t)| \ll |R_{AV}(t)|$, that is, the multisensory response is captured well by the mode decomposition in equation 6.7 then the corresponding time dependent amplitudes, $\xi_{AV_A}^i(t)$ and $\xi_{AV_V}^j(t)$, must be the activity of the auditory and the visual sources respectively, in response to multisensory stimulation. In contrast, the time dependent amplitudes $\xi_A^i(t)$ and $\xi_V^j(t)$ represent the response of the auditory and visual subsystems to unisensory stimulation in their respective modality. In order to determine the temporal influences of multisensory stimulation on the unisensory subsystems, the relevant variables to consider are the differ-

ences $(\xi_{AV_A}^i(t) - \xi_A^i(t))$ and $(\xi_{AV_A}^j(t) - \xi_V^j(t))$ for audition and vision respectively. The first two terms on the right hand side of equation 6.7 describe the influence of the auditory and visual subsystems on the multisensory response. The remainder, $R_{AV_{spatial}}(t)$ must therefore be a measure of the influence of any additional sources that are unique to multisensory stimulation. The subscript ‘spatial’ indicates that the remainder is due to additional sources and not due to the temporal modulation of the unisensory subsystems.

6.4.3 Adjoint basis decomposition.

Now we must address certain practical issues regarding the application of the above methodology to experimental data. The main issues are i) The choice of spatial modes ii) Extracting the temporal behavior of these modes for the multisensory condition. In addressing the first issue it is important to note that spatial mode decompositions are not unique and biased towards certain criteria, such as the minimization of the square error (principal component analysis) or statistical independence (independent component analysis (Jung et al., 2001)). Other techniques use apriori knowledge about the system’s dynamics to determine interpretable modes (Tesche et al., 1995). In our analysis we used the principal components to obtain a mode decomposition of the spatio-temporal data. In order to address the second issue we used an adjoint basis decomposition to obtain the time dependent amplitudes $\xi_{AV_A}(t)$ and $\xi_{AV_V}(t)$. In general the modes ϕ_A^i and ϕ_V^j will not be orthogonal. Therefore the time dependent amplitudes cannot be calculated by projecting the multisensory signal onto the unisensory modes but must be obtained by a projection onto the corresponding adjoint vectors in dual space. The adjoint vectors have the following properties,

$$\phi_A^i \cdot \phi_A^{j\dagger} = \delta_{ij}, \quad \phi_V^i \cdot \phi_V^{j\dagger} = \delta_{ij}, \quad \text{and}, \quad \phi_A^i \cdot \phi_V^{j\dagger} = \phi_V^i \cdot \phi_A^{j\dagger} = 0 \quad (6.8)$$

The dimensions of the basis space is in general larger than the total number of modes used to describe the spatiotemporal evolution, that is, $k, k' \leq n$. Therefore, the adjoint basis is not defined uniquely. We must make the additional assumption that the adjoint basis may be defined in the space spanned by the original basis vectors,

$$\phi_A^{i\dagger} = \sum_{j=1}^k \lambda_A^j \phi_A^j + \sum_{j=1}^{k'} \lambda_V^j \phi_V^j \quad (6.9)$$

We may similarly define $\phi_V^{i\dagger}$.

The time dependent amplitudes may be obtained by projecting the multisensory data onto the appropriate adjoint mode.

$$\xi_{AV_A}^i(t) = AV(t) \cdot \phi_A^{i\dagger}, \text{ and, } \xi_{AV_V}^i(t) = AV(t) \cdot \phi_V^i \quad (6.10)$$

6.5 Experimental results and discussion.

Here, we apply the methods developed in section 6.4 to our EEG experiment. In section 6.5.1 we will first express the data obtained using the unisensory control conditions in terms of spatial modes and its corresponding time dependent amplitudes. These modes allow us to construct a basis that may then be used to describe the scalp potentials measured in response to multisensory stimulation. The time dependent amplitudes will be obtained using an adjoint basis decomposition. In section 6.5.2 we compare our method with the results obtained using the summative model. This is particularly important since the applicability of our method relies on the relationship $|R_{AV_{spatial}}(t)| \ll |R_{AV}(t)|$. We demonstrate that this relationship is indeed true across all subjects and experimental conditions. One of our main goals was to examine the temporal modulations of the unisensory components in response to multisensory stimulation. In section 6.5.3 we will address this specific issue. Finally we will examine the contributions due to uniquely multisensory sources in section 6.5.4.

First, a note on the terminology that will be used throughout the remainder of this chapter. The modes obtained due to auditory-only and visual-only stimulation were denoted as ϕ_A^i and ϕ_V^j respectively. These modes are associated with sources responsible for the scalp potentials measured in response to unisensory stimulation and will be referred to as auditory sources, visual sources or unisensory sources when appropriate. The temporal evolution of the unisensory sources in response to auditory-only and visual-only stimulation was denoted

by $\xi_A^i(t)$ and $\xi_V^j(t)$. These will be referred to as the unisensory time series. The temporal evolution of the unisensory sources due to multisensory stimulation, $\xi_{AV_A}^i(t)$, and $\xi_{AV_V}^j(t)$ will be referred to as the multisensory-A and the multisensory-V time series (Compare equations 6.4 and 6.5 with 6.10).

We will primarily be interested in two residual time series,

1. $R_{AV_{temporal}}(t) = \sum_{i=1}^k (\xi_{AV_A}^i(t) - \xi_A^i(t))\phi_A^i - \sum_{j=1}^k (\xi_{AV_V}^j(t) - \xi_V^j(t))\phi_V^j$. This time series gives us an indication of how divergent the multisensory-A and multisensory-V time series are from the respective unisensory time series, that is, how does multisensory processing modulate the activity of the unisensory sources.
2. $R_{AV_{spatial}}(t) = AV(t) - (\sum_{i=1}^k \xi_{AV_A}^i(t)\phi_A^i + \sum_{j=1}^{k'} \xi_{AV_V}^j(t)\phi_V^j)$. This time series is an indicator of the activity of sources other than the unisensory sources that are activated in response to multisensory stimulation.

6.5.1 Mode decomposition

We first decompose the measured potentials, $A(t)$ and $V(t)$, due to auditory-only and visual-only stimulation respectively, in terms of spatial modes and their corresponding time varying amplitudes.

$$A(t) = \xi_A(t)\phi_A + R_A(t) \quad \text{and} \quad V(t) = \xi_V^1(t)\phi_V^1 + \xi_V^2(t)\phi_V^2 + R_V(t)$$

Figure 6.3 shows the first auditory principal component and the first two visual principal components for a particular subject. The topography of the auditory mode is characterized by a frontal-central maximum that corresponds closely to the N1 component of the auditory event related potential and has been attributed to two bilaterally symmetric sources located along the primary auditory cortex. Across subjects, the first auditory mode explains 85-96% of the observed variance. The two prominent visual modes are characterized by the presence of an occipital peak and account for 85-95% of the variance. The modes ϕ_A , ϕ_V^1 , and ϕ_V^2 form the basis space in which we will examine the responses to multisensory stimulation. The multisensory time series are obtained by projecting the potentials measured on the scalp

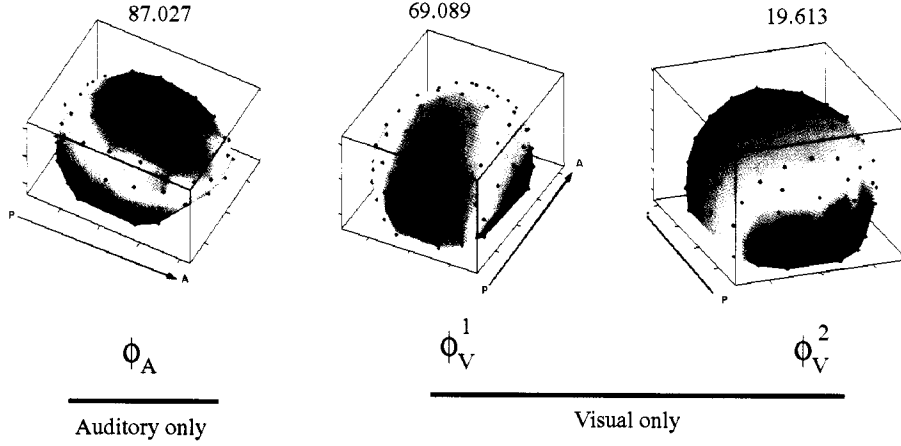


Figure 6.3: Principal components corresponding to auditory-only and visual-only stimulation for one subject. The number above each figure represents the percentage of variance explained by the corresponding mode. The arrows adjacent to each figure point in the anterior-posterior direction marked as A and P respectively.

in response to multisensory stimulation, $AV(t)$, onto the adjoint basis vectors ϕ_A^\dagger , $\phi_V^{\dagger 1}$ and $\phi_V^{\dagger 2}$.

6.5.2 Comparison between the summative model and the mode decomposition.

As a first step we wish to quantify the contributions due to the auditory-only sources, the visual-only sources and uniquely multisensory sources. Here, it is important to identify the difference between our proposed method and the summative model that has traditionally been used in similar studies.

The summative model may be expressed as follows

$$\mathbf{AV}(t) = \mathbf{A}(t + \tau_1) + \mathbf{V}(t + \tau_2) + \mathbf{R}_{AV}(t) \quad (6.11)$$

where, $\mathbf{AV}(t)$ are the potentials measured on the scalp in response to multisensory stimulation. In our experiment we have a number of conditions where the inter-onset interval between two stimuli of a pair are separated by either $dT=100\text{ms}$ or $dT=200\text{ms}$. Consider the case where the auditory stimulus is followed by a visual stimulus 200ms later. For

this example the auditory ERP must be shifted by $\tau_1 = 0\text{ms}$ and the visual ERP must be shifted by $\tau_2 = 200\text{ms}$ in order to compare the multisensory response with the summed unisensory response. The residual time series, $\mathbf{R}_{\mathbf{AV}}(t)$, resulting from the summative model contains contributions from two different factors, the temporal modulation of the activity of unisensory sources due to multisensory stimulation and additional sources of activation other than the unisensory sources. A reconstruction of the measured potentials obtained by projecting $\mathbf{AV}(t)$ onto the auditory and visual modes is given by

$$\mathbf{AV}(t) = \xi_{AV_A}^1(t)\phi_{\mathbf{A}} + \xi_{AV_V}^1(t)\phi_{\mathbf{V}}^1 + \xi_{AV_V}^2(t)\phi_{\mathbf{V}}^2 + \mathbf{R}_{\mathbf{AV}_{\text{spatial}}}(t) \quad (6.12)$$

In order to quantify the contributions due to the unisensory sources we define a goodness of fit variable using the residual time series, $\mathbf{R}_{\mathbf{AV}_{\text{spatial}}}(t)$, as follows

$$G_{\text{spatial}} = 1 - \frac{1}{\tau} \int_0^T \mathbf{R}_{\mathbf{AV}_{\text{spatial}}}(t)^2 dt$$

Where, $\tau = \int_0^T AV(t)^2 dt$, is a normalization constant and T is the length of the trial. The quantity E_{spatial} is a measure of how much of the observed multisensory data may be explained by the activity of unisensory sources alone. $G_{\text{spatial}} \in [0, 1]$. If $G_{\text{spatial}} = 1$, then $AV(t)$ may be ascribed entirely to the activity of unisensory sources. Similarly we define a goodness of fit for the summative model as follows,

$$G_{\text{summative}} = 1 - \frac{1}{\tau} \int_0^T \mathbf{R}_{\mathbf{AV}}(t)^2 dt$$

Where $G_{\text{summative}} \in [0, 1]$. If $G_{\text{summative}} = 1$, then we may assume that the temporal evolution of the unisensory sources in response to unisensory stimulation is identical with its evolution in response to multisensory stimulation. In addition we may also assume that there are no additional sources of activity unique to multisensory stimulation, that is, there is no measured interaction between the unisensory subsystems. Figure 6.4 is a comparison between G_{spatial} and $G_{\text{summative}}$. The values indicate the mean goodness of fit across subjects and the standard error is plotted over the bars. Each pair of vertical bars represent the goodness of fit for a particular experimental condition characterized by

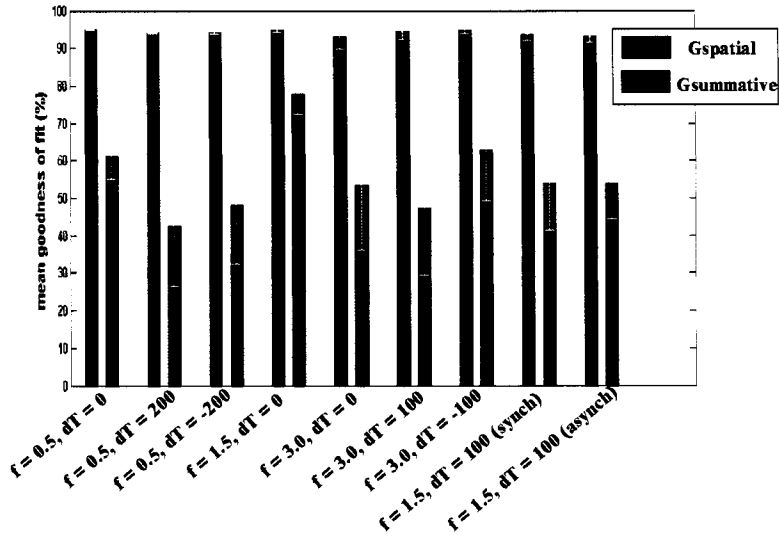


Figure 6.4: Shown is a comparison between $G_{spatial}$ and $G_{summative}$ across subjects for each experimental condition.

a point in the $dT - ISI$ space. The values of $G_{spatial}$ are larger than those of $G_{summative}$ across subjects and experimental conditions indicating that the changes in the topography of potentials measured on the scalp in response to multisensory stimulation may be attributed primarily to the temporal modulation of the unisensory sources. The measured potential modulations due to uniquely multisensory sources account for less than 10% of the variance seen in the data.

Our analysis is based entirely on the ERP. It is important to note here that there may be additional sources that are insensitive to our measure due to their location in the brain, the variability in their response to a repeated stimulus or their low amplitude. Therefore, we do not imply here that the unisensory sources are the primary loci of multisensory processing. Figure 6.5 compares the difference between the reconstructed scalp potentials obtained using an adjoint basis decomposition with that obtained using the summative model. The traces are shown for two channels, CZ (central) and OZ (occipital). Based on the results shown in Figures 6.4 and 6.5 we can conclude that $|\mathbf{R}_{AV_{spatial}}(t)| \ll |\mathbf{R}_{AV}(t)|$, thus justifying the use of our method to further analyze the temporal features of the unisensory subsystems.

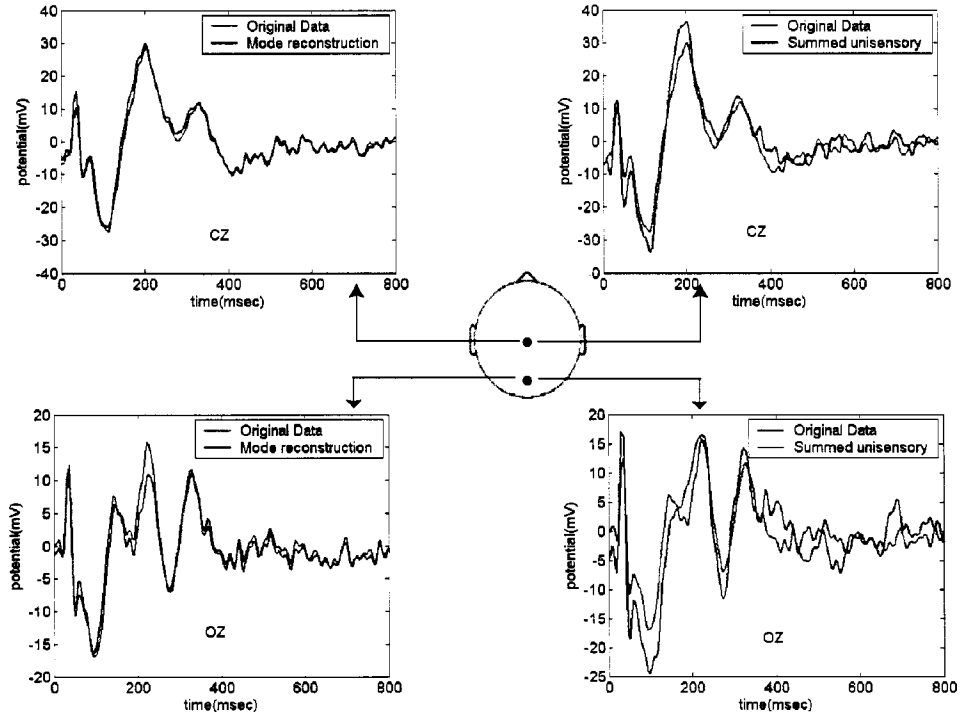


Figure 6.5: Comparison between the summative model (right column) and the time series obtained by projecting $AV(t)$ onto the modes ϕ_A and ϕ_V (left column) and reconstructing the time evolution $AV(t)$.

6.5.3 Early auditory-visual interactions.

We now examine the multisensory-A and the multisensory-V time courses in order to determine the latency of interactions between the corresponding sources. It has been suggested in the literature that the interaction between the auditory and the visual sources occurs at an early latency in the vicinity of 40ms. This view is a radical departure from the idea that multisensory interactions occur in higher processing areas following extensive unisensory processing. In order to determine the interaction between the auditory and the visual sources we consider the residual time series $R_{AV_{temporal}}$ which is an indicator of the temporal modulations of the unisensory sources. The topographs in Figure 6.6 are the first two principal components obtained from the time series $R_{AV_{temporal}}(t)$ for latencies between 30-90ms.

We notice an early onset of a strong visual component characterized by an occipital peak in the residual topographs. The first principal component, across subjects, explains between 90%-96% of the variance in the latencies between 30-90ms indicating that the temporal evolution of the visual sources is strongly affected by a simultaneously presented auditory stimulus and may be represented by the time dependent amplitude of the first principal component of the residual. The residual topographs closely resemble the second principal component ϕ_V^2 obtained using visual only stimulation. The latency of the first peak of this occipital component extends over a wide range across subjects, from approximately 40ms to 60ms.

6.5.4 Uniquely multisensory sources.

6.5.4.1 The Drift regime.

In section 6.5.2 we showed that $|R_{AV_{spatial}}(t)| \ll |R_{AV}(t)|$, implying that most of the effects of multisensory processing, that are obtained using the measured scalp potentials may be attributed to sources that are also activated in response to unisensory stimulation in each modality. The residual signal that remains $R_{AV_{spatial}}(t)$, therefore, must contain the influences due to uniquely multisensory sources. The peaks in this residual time series correspond to the latencies where these sources are maximally activated. The power associated with

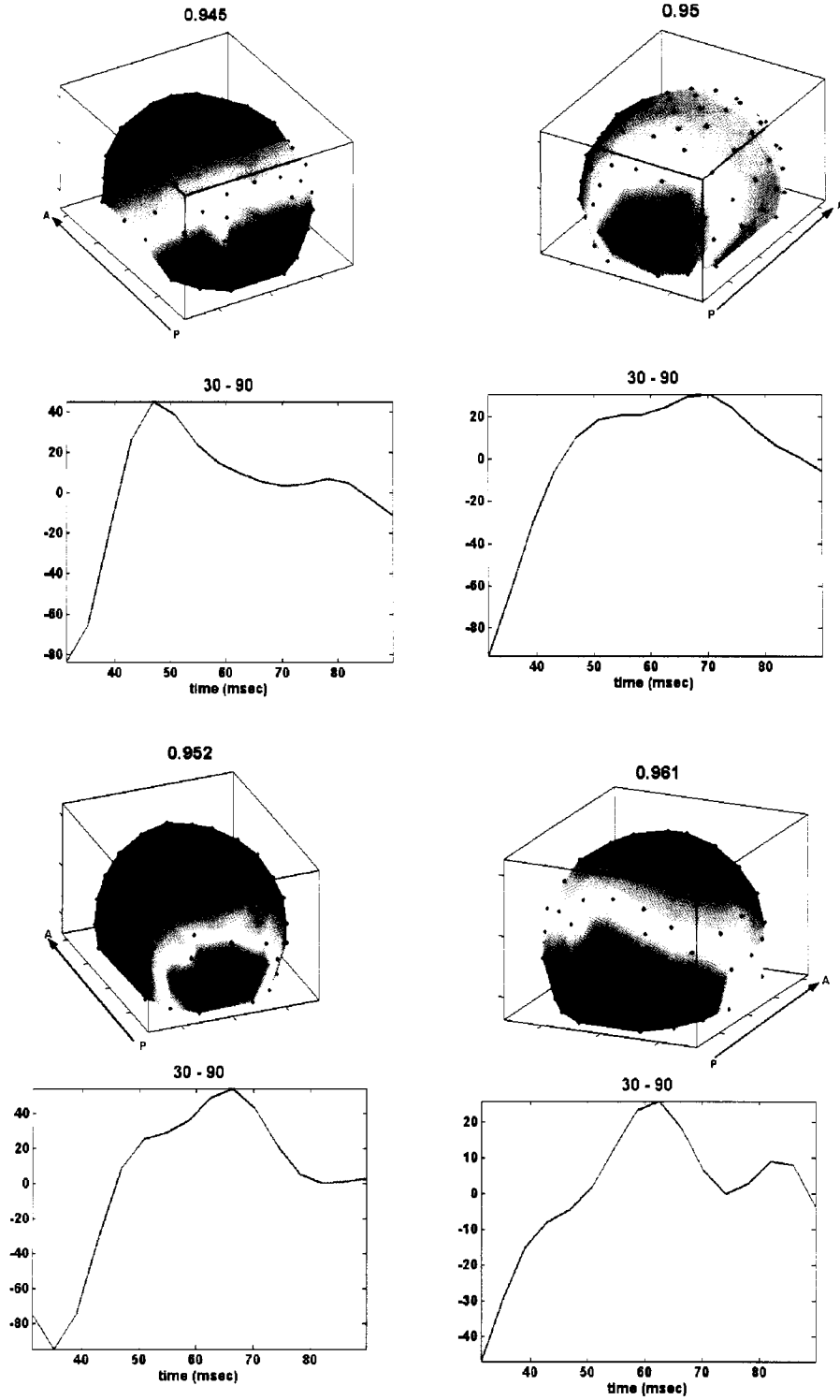


Figure 6.6: First principal component of the residual time series $R_{AV_{temporal}}$. Each topograph corresponds to one subject. The temporal evolution of the components are shown below the respective topographs. The normalized eigenvalues are shown indicated above each component.

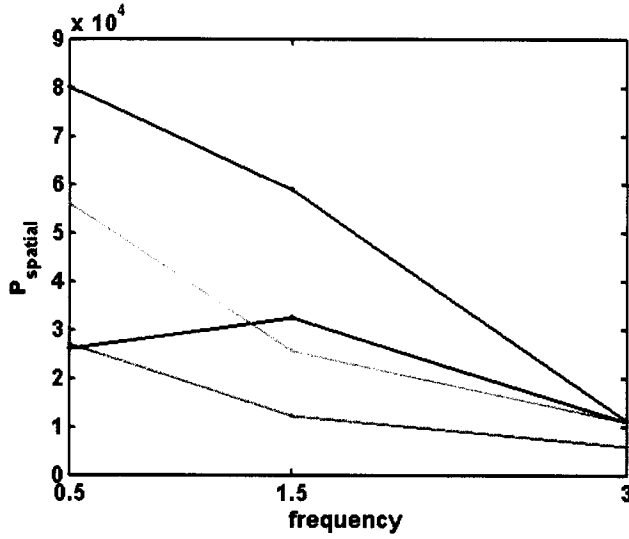


Figure 6.7: Shown is $P_{spatial}$ as a function of frequency. Each color represents a different subject.

$R_{AV_{spatial}}(t)$ may be defined as follows

$$P_{spatial} = \int_0^T \sum_{electrodes} R_{AV_{spatial}}(t)^2 dt \quad (6.13)$$

Here $T = 335$ ms, is the duration of an epoch when the stimulation rate is 3Hz. The power $P_{spatial}$ may be attributed to additional sources that are activated in response to multisensory stimulation. In Figure 6.7 $P_{spatial}$ is shown as a function of the three frequencies used in our experiment.

Each trace corresponds to a particular subject. Three out of four subjects show a monotonic decrease in the value of $P_{spatial}$ as a function of frequency. The smallest value of $P_{spatial}$ corresponds to a stimulation rate of 3Hz across all subjects. $P_{spatial}$ is the power associated with uniquely multisensory sources. A decrease in the value of $P_{spatial}$ as a function of increasing frequency implies that either the activity or the number of these additional sources decreases as a function of frequency. A frequency of 3Hz corresponds to the drift regime where subjects' perception of the stimulus varied across the length of a trial. Our earlier fMRI study (see Chapter 4) indicated that the perception of drift is associated with the deactivation of the inferior parietal areas and the posterior midbrain leaving only a network consisting of the primary sensory and prefrontal areas. Figure 6.7 shows that the power

decreases as a function of increasing frequency with the lowest value being associated with a frequency of 3Hz. A decrease in power may be attributed to two factors, one being a decrease in the activation of the uniquely multisensory sources and a second being a decrease in the number of sources. Our results in conjunction with the results of the earlier fMRI experiment suggest that the decrease is due to deactivation of some of the sources involved in multisensory processing.

6.5.4.2 Synchrony vs Asynchrony.

Now consider the experimental condition represented by the red box in Figure 6.2. Here, the stimuli were presented at a rate of 1.5Hz with an interonset interval of $dT = 100\text{ms}$. We established in earlier behavioral experiments that at this point in the dT -ISI parameter space the perception of the stimulus is bistable. Subjects' perception of the stimulus depended on the instructions. Two sets of instructions were given and the scalp potentials monitored for each case, while the stimulus conditions remained the same. In the first condition, AV_{synch} , subjects were instructed to try to perceive the stimuli as synchronous events, while in the second condition, AV_{asynch} , subjects were instructed to try to perceive the stimulus as sequential, that is, auditory precedes visual. The residual time series, $R_{AV_{spatial}}(t)$ was computed for each of these conditions and the difference between the two is denoted by $R_{SynchAsynch}(t)$. Since the stimulus in both conditions is exactly the same, $R_{AV_{SynchAsynch}}(t)$ is a good indicator of the differences between the percepts of synchrony and asynchrony. The time series $R_{AV_{SynchAsynch}}(t)$ shows the presence of two broad peaks that indicate the latencies at which the uniquely multisensory sources are maximally active. The latency of the first peak extends from approximately 80ms to 150ms while the second peak extends from approximately 160ms to 300ms. The principal components associated with each peak are shown in Figure 6.8. The topographs show a distinct bi-lobed occipital component which suggests the presence of an inferior parietal source of activation in addition to a prefrontal source activation. A notable result of the earlier fMRI study is that the perception of asynchrony activates a network consisting of auditory, visual, prefrontal and parietal areas, in particular the left inferior parietal lobule. The perception of synchrony is associated with the disengagement of the inferior parietal lobule. Similar results have been shown by

Bushara et al. (2001) who suggest a role of the IPL in asynchrony detection. Our results, based on the topographs and the corresponding time series in Figure 6.8, suggest that an inferior parietal source with peak activity near 200 ms may explain the main features of the residual time series, $R_{SynchAsynch}$.

6.6 Summary

In this chapter we have presented a method that allows us to distinguish the various contributions to multisensory integration and demonstrated that the early effects captured using scalp EEG may be primarily attributed to the unisensory subsystems. These early modulations occur at a latency of approximately 40ms indicating that multisensory processing has a significant feedforward component. This result adds to the current hypotheses in the literature and suggesting early auditory-visual interactions and is supported by evidence in the electrophysiological and neurophysiological literature on multisensory integration. One of the main contributions of our work is methodological and allows us to examine, in a principled manner, the summative model previously proposed in the literature. The asymmetric occipito-temporal components corresponding to the late stages of multisensory processing (~150-300ms) post-stimulus suggest possible generators in the inferior parietal cortex which has been demonstrated as a significant multisensory convergence zone in previous of experiments.

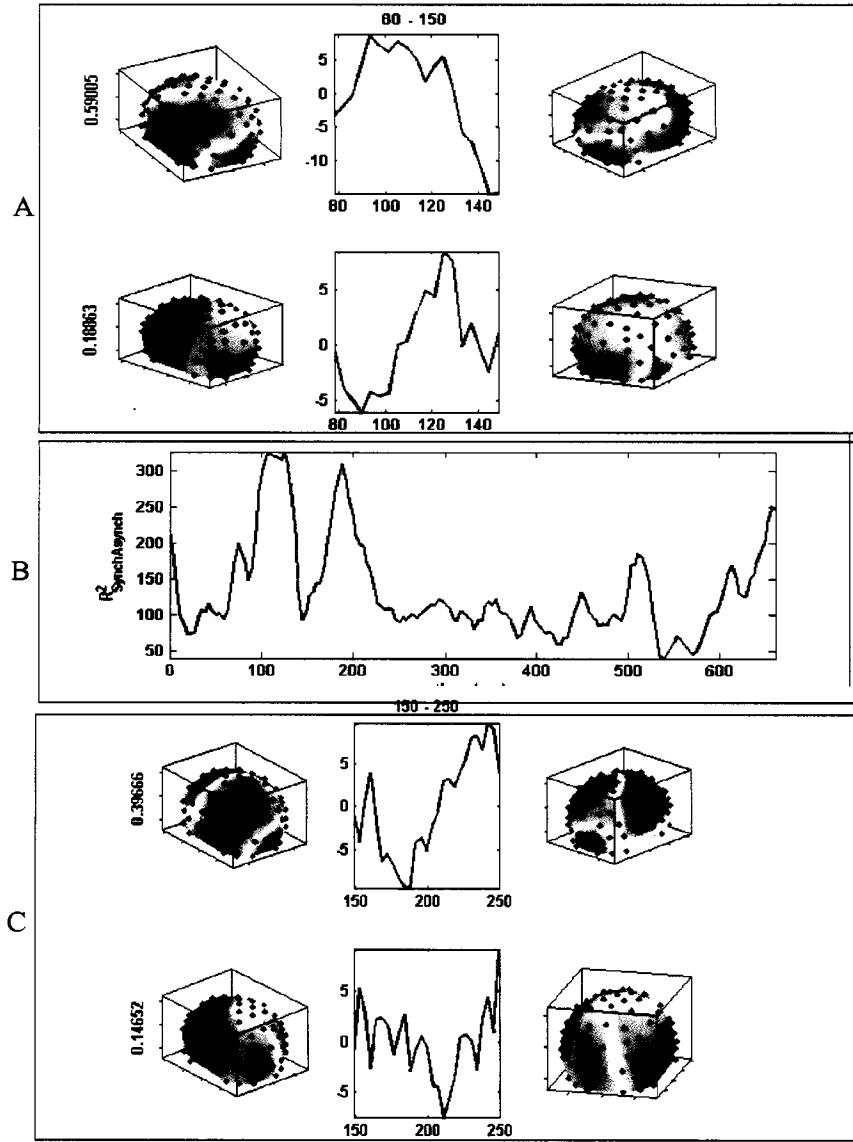


Figure 6.8: Two different sets of topographs corresponding to the residual signal $R^2_{SynchAsynch}$ is represented by the black trace in panel C of the figure. The topographs in panel A are associated with the first peak of $R^2_{SynchAsynch}$ extending from a latency of 80ms to 150ms. The top two topographs in panel A show different views of the first principal component of $R_{SynchAsynch}$ for the first peak. The bottom two topographs in panel A show different views of the second principal component. The two traces in panel A show the time course of the respective principal components. In panel C the principal components associated with the second peak are shown. The normalized eigenvalues of each component are shown adjacent to the respective component.

Chapter 7

Summary

In chapter 1 we posed the following questions,

1. What is a mathematically tractable approximation of the functional units?
2. How does the dynamics of functional units manifest itself in behavior and perception?

In response to the first question we provided a general description to obtain a low dimensional description of globally coupled dynamical systems. Our approach advances the current understanding of the dynamics of large ensembles of globally coupled dynamical systems in two respects. One, we demonstrated, that the dynamics of globally coupled dynamical systems with parameter dispersion may be treated as a problem of spatiotemporal pattern formation with the space being defined by the dispersed parameter. An advantage of reformulating the problem in terms of spatiotemporal pattern formation is that an extensive mathematical machinery developed to deal with similar problems becomes available to us (Cross and Hohenberg, 1993). Our solution to the problem involved a spatiotemporal mode decomposition that allowed a low dimensional description. In general, such a decomposition is always possible. Whether the decomposition provides a good description of the network dynamics is dependent on the specific instantiation of the network. A limitation of our methodology is that it does not provide an apriori ‘ideal’ decomposition. A second contribution of the approach detailed in chapter 2 is related to the specific implementation of globally coupled Fitzhugh-Nagumo neurons (Fitzhugh, 1961; Nagumo et al., 1962). We demonstrated that

the behavior of this system may be adequately described by a four dimensional system that resembles two coupled Fitzhugh-Nagumo neurons. This low dimensional system of coupled excitable units shows synchronous and phase clustered oscillations in addition to oscillator death as a function of coupling strength and parameter dispersion that closely mimics the behavior of the network.

In addressing the second question we examined two seemingly diverse phenomena, sensorimotor coordination and multisensory integration. We devised a set of experiments that explore the commonality between coordinative behavior and multisensory perception. The former case was conceived as a driven functional unit while the latter as two interacting functional units. While a vast literature precedes our modelling and experimental work in sensorimotor coordination (Kelso, 1995), multisensory perception seen from the context of dynamical systems theory does not share a similar precedent. One of our main contributions in this area is to provide a new framework within which the phenomenon of multisensory integration may be viewed. Second, we provide an empirically motivated architecture of the multisensory network that tries to capture the qualitative features of the experiment. The form of the coupling between the nodes of the network leads to a characteristic dynamics of the functional units that may then interpreted in terms of the perception of a multisensory stimulus sequence. Asymmetries in perception were modelled using asymmetries in temporal delays and intrinsic time scales of the functional units involved.

Our experimental and theoretical results on sensorimotor coordination (chapter 2) demonstrate that the coupling between the external environment and the rhythmic movement of a finger must be minimally parametrically driven. In addition we make a prediction that there are two modes of phase transition from an anti-phase to an in-phase mode of coordination, one mediated by a change in the amplitude of finger movement the other mediated by a shift in phase. This prediction is now open to experimental verification (DeGuzman et al., 2002).

We further probed the neural correlates of multisensory integration in our fMRI experiment (chapter 5). Our findings show that different distributed networks are activated when perceptually synchronous and asynchronous stimuli are processed. In particular, the superior colliculus is associated with the perception of synchrony and the inferior parietal cortex

with the perception of asynchrony of auditory-visual signals. However, when there is no clear percept, these network components are disengaged and only a residual network comprised of sensory and frontal areas remains active. These results demonstrate that the processes of perceptual integration and segregation engage and disengage different brain subnetworks, but leave only a rudimentary network in place if no clear percept is formed.

Our EEG experiments contribute to the current knowledge from a methodological perspective. We derive a method that allows us to highlight different contributions to multisensory integration, that due to the interaction of the unisensory subsystems and due to additional sources that are activated only in response to multisensory stimulation. The current approach to examine the nature of multisensory interactions has been restricted to the summative model (Giard and Peronnet, 1999). In chapter 6, we demonstrate some of the deficiencies of the summative model and compare it with our mode decomposition approach. We demonstrate that over 90% of the measured scalp potentials in response to multisensory stimulation may be attributed to the interaction of unisensory subsystems. The latency at which these interactions begin to take effect are identified to be in the vicinity of 30ms suggesting that auditory-visual interactions occur remarkably early in time. Whether this interaction is mediated by direct anatomical connections (Falchier et al., 2002; Rockland and Ojima, 2003) between the unisensory subsystems is an open question. However, our investigations in addition to results from a number of recent studies suggest an additional layer of interactions in the sensory processing hierarchy (Foxe and Schroeder, 2005). Evidence of early interactions challenges some of the traditional views in Neuroscience where the unisensory subsystems are assumed to retain their sensory fidelity until later stages of processing (Mesulam, 1998). Our results also find support in some recent work suggesting the existence of early interactions between the unisensory subsystems (Giard and Peronnet, 1999; Molholm et al., 2002).

Appendix A

fMRI materials and methods

A.1 Subjects

Eighteen subjects, aged between 25 to 37, participated in the behavioral experiment. Thirteen of those also participated in the fMRI experiment. All the subjects were in good health with no past history of psychiatric or neurological diseases. Informed consent was collected from each subject prior to the experiment, and the study was approved by the Florida Atlantic University Institutional Review Board.

A.2 Behavioral Experiment

A.2.1 Experimental design

The behavioral experiment consisted of two sessions: a training session with 5 conditions and a session with 36 conditions, each in a 30-second block. The participants were presented auditory and visual stimuli (440 Hz-30 ms tones from a speaker placed in front at a distance of 2.5 meters and 30 ms flashes from a red LED) at different timing parameters (intra-stimulus interval ranging from -200 to 200 and stimulation rate from 0.5 to 3.5 Hz) and asked to perceive these stimuli as a synchronized auditory visual event. The participants were asked to categorize the percept as ‘simultaneous’ (S) if they had perceived tone and flash as synchronized events throughout the run. If the perception of synchrony was not established, then the subjects were asked to classify their percept according to the following

categories: ‘AV (auditory before visual stimuli)’, ‘VA’, or ‘Can’t tell’. The last category was described by the participants as equivalent to phenomenon of phase drift (“the stimuli seemed to be together, then they drifted and got back together again”) or streaming (“the two stimuli were in two separate streams with no reportable relationship to each other”).

A.2.2 Behavioral Performance

Only those subjects with robust percepts, determined in the training session, participated in the main study. For example, two subjects who reported perception of synchrony throughout the entire timing parameter space were not included in the final behavioral performance analysis. A percept was classified as robust, if the percept remained the same throughout the entire block of 30 seconds after some initial transient time. At the end of each block, the subjects were asked to report their percept. The perceptual strength in the timing parameter space was quantified by the number of subjects reporting a given percept.

A.3 fMRI Experiment

A.3.1 Experimental design

The tasks in fMRI experiments consisted of three functional runs of sensory (auditory, visual and auditory-visual) stimulation, and rest conditions in on-off block design. In the stimulation conditions, a series of 440 Hz-30 millisecond auditory tones or/and 30 millisecond visual red flashes were presented in a pair of goggles and earphones in the scanner. The first and second runs were each 27.3 minutes long with 24 second-on and 18 second-off block repeated 3 times per condition. In the first and the second runs, there were 6 unimodal conditions and 7 bimodal conditions for the stimulation rates of 0.5, 1.5 and 3.0 Hz and the time intervals $(-200, 0, 200)$, $0, (-100, 0, 100)$ milliseconds between the onsets of a pair of an auditory tone and a visual flash. Subjects were instructed to try to perceive a pair of auditory and visual stimuli as simultaneous events in order to maintain the same level of attention towards either modality. The third run was 9.0 minutes long with 3-second -visual instruction, 24 second-on and 18 second-off block repeated 6 times per condition. In this run, there were 2 bimodal conditions with 100 milliseconds time interval between the onsets of

auditory and visual stimuli presented at 1.0 Hz. The subjects were instructed with a visual cue at the beginning either to perceive the auditory and visual stimuli simultaneously or sequentially. The conditions in each run were presented in random order. Before the actual experiment, each subject was taken to the behavioral lab to familiarize them with the tasks. All stimuli presentations were performed with the software, Presentation (Neurobehavioral Systems, Inc., San Francisco, CA).

A.3.2 Image acquisition

A 1.5 Telsa GE Signa scanner was used to acquire T1-weighted structural images and functional EPI images for the measurement of the blood oxygenation level-dependent (BOLD) effect (Kwong et al., 1992; Ogawa et al., 1992). The acquisition scheme and parameters used for the functional scans (546 scans in the first and second sessions and 180 scans in the third session) are as follows: echo-planar imaging, gradient recalled echo, TR = 3000 ms, TE = 40 ms, flip angle = 90 degree, 64 x 64 matrix, 30 axial slices each of thickness 5 mm acquired parallel to anterior-posterior commissural line.

A.3.3 Data Analysis

The data were preprocessed and analyzed using Statistical Parametric Mapping (SPM2; Wellcome Department of Cognitive Neurology, London, UK) (Friston et al., 1995, 1999). Motion correction to the first functional scan was performed within subject using a six-parameter rigid-body transformation. All the 13 subjects included in this analysis had less than 4mm of translation in all directions and less than 2.0 degrees of rotation about the three axes. The mean of the motion-corrected images was then coregistered to the individual's 30-slices structural image using a 12-parameter affine transformation. The images were then spatially normalized to the Montreal Neurological Institute (MNI) template by applying a 12-parameter affine transformation, followed by a nonlinear warping using basis functions (Ashburner and Friston, 1999). These normalized images were interpolated to 2mm isotropic voxels and subsequently smoothed with a 4 mm isotropic Gaussian kernel. A random-effects, model-based, statistical analysis was performed with SPM2 (Friston et al., 1995, 1999) in a two level procedure. At the first level, a separate general linear model of the form:

$Y = X\beta + \epsilon$, was specified for each subject, where $X = [b \ X_1 \ X_2 \ \dots]$, $\beta = [1 \ \beta_1 \ \beta_2 \ \dots]^T$, and $\epsilon = N(0, \sigma^2)$, b is a constant. Each $X_1, X_2, \dots, X_i \dots$ consists of series of 0 for off-blocks and 1 for on-blocks, and represent different stimulation conditions in each functional run and 6 motion parameters obtained from the realignment. Thus, $\beta = (X^*{}^T X^*)^{-1} X^*{}^T KY$, where $X^* = KX$, and K is a matrix constructed with the SPM2 standard hemodynamic response function. Individual contrast images were created by correlating the brain response with the aforementioned covariates (X_i 's) for each subject. Global differences among subjects were controlled by proportional scaling. The individual contrast images were then entered into a second-level analysis, using a separate one-sample t-test for each term in the general linear model. We then thresholded these summary statistical maps at $p < 0.001$ (uncorrected for multiple comparisons). These maps were overlaid on a high-resolution structural image in the Montreal Neurological Institute (MNI) orientation. We used these summary maps to perform a region of interest (ROI) analysis on different conditions. We used Analysis of Functional Neural Images (AFNI) software (Cox, 1996) to extract the average time series from the activated regions corresponding to the on-off blocks of different conditions.

A.3.4 Tables and legends

The following table displays significantly activated brain areas, the anatomical location in MNI coordinates, cluster size ($k \geq 10$ voxels), T and z-scores at $p < 0.001$ (uncorrected) for all the bimodal conditions. The activations associated with negative contrasts are also listed with a negative sign in conditions. The following abbreviations are used in this table: M1 (modality 1), M2 (modality 2), A (auditory), V (visual), L (left), R(right), G (gyrus), SFG (superior frontal gyrus), STS (superior temporal sulcus), MOG (middle occipital gyrus), MFG (middle frontal gyrus), MeFG (medial frontal gyrus), PHG (parahippocampal gyrus), TTG (transverse temporal gyrus), IPL (inferior parietal lobule), and PM (posterior mid-brain).

Conditions M_1 M_2 (Δt ms, f Hz)	Brain areas	MNI coordinates (x, y, z)	Cluster size	T (z-score)
AV (0, 0.5)	R SFG	(18, 36, 40)	30	6.9 (4.31)
	R SFG	(-12, 32, 50)	19	6.24 (4.09)
	L STS	(-58, -44, 6)	21	8.24 (4.69)
	R MOG	(-38, -62, 0)	11	6.27 (4.10)
-AV (0, 0.5)	R Cuneus	(2, -78, 34)	16	5.19 (3.69)
AV (0, 1.5)	L MFG	(-50, 2, 42)	33	6.05 (4.02)
	L MFG	(-6, -4, 60)	27	5.6 (3.85)
	L STS	(-54, -34, 8)	255	7.57 (4.51)
	R STS	(64, -40, 12)	43	7.57 (4.51)
	L MOG	(-38, -66, -2)	16	5.13 (3.66)
- AV (0, 1.5)	L. PHG	(-18, -52, -10)	116	9.84 (5.06)

(a)

AV (100, 1.0) ("Synchronize")	L. MFG	(-38, 46, -8)	50	13.35 (5.12)
	R. STG	(46, -32, 10)	20	6.43 (3.85)
	R. MFG	(48, 50, -4)	21	5.48 (3.55)
- AV (100, 1.0)	R. PM	(0, -30, -12)	40	10.93 (4.79)
	L. Cuneus	(-4, -62, 8)	12	8.83 (4.42)
VA (-200, 0.5)	R. MFG	(50, 16, 36)	96	6.48 (4.17)
	R. STG	(62, -36, 10)	21	6.41 (4.15)
	L. TTG	(-34, -52, 46)	22	7.12 (4.37)
	L IPL	(-34, -34, 10)	107	5.78 (3.92)
	R. MTG	(60, -34, -2)	94	6.71 (4.25)
	R. MOG	(30, -84, 2)	13	6.12 (4.05)
AV (200, 0.5)	R. MFG	(36, 54, 2)	60	5.05 (3.63)
	L IPL	(-40, -50, 40)	28	4.66 (3.46)
	L MOG	(-34, -82, -2)	13	4.79 (3.51)
- AV (200, 0.5)	R STG	(60, -2, 0)	13	7.86 (4.59)
	L. Cuneus	(-12, -64, 6)	12	6.0 (4.0)

(b)

AV (100, 1.0) ("Perceive AV")	R. STG	(56, -16, 2)	60	10.57 (4.73)
	L. IPL	(-36, -56, 46)	95	16.37 (5.44)
	L. MFG	(-46, 18, 34)	13	7.08 (4.02)
AV (0, 3.0)	L. STG	(-46, -20, 6)	274	6.57 (4.20)
	R. STG	(46, -24, 6)	45	6.57 (4.20)
	R. STG	(58, -28, 4)	36	8.38 (4.72)
	L. IOG	(-38, -70, -6)	33	7.30 (4.43)
VA (100, 3.0)	R. IFG	(56, 12, 30)	64	6.4 (4.15)
	L STG	(-48, -16, 0)	40	7.8 (4.57)
	R STS	(62, -34, 6)	32	7.52 (4.49)
	R MTG	(50, -64, -4)	40	6.48 (4.17)
	L ITG	(-48, -68, -2)	36	6.91 (4.31)
	R MOG	(38, -82, -6)	23	7.63 (4.52)
-VA (100, 3.0)	R. Cuneus	(4, -82, 10)	58	8.09 (4.65)
AV (100, 3.0)	R IFG	(56, 8, 20)	13	4.41 (3.34)
	L IFG	(-56, 12, 20)	16	5.61 (3.86)
	R STG	(50, -12, -2)	46	6.22 (4.08)
	L STG	(-48, -18, 2)	167	5.41 (3.78)
	R MTG	(42, -58, -2)	98	7.61 (4.52)
	R MOG	(28, -88, -4)	124	8.84 (4.83)

Figure A.1: Significantly activated brain areas. The anatomical location in MNI coordinates, cluster size ($k \geq 10$ voxels), t and z scores at $p < 0.001$ (uncorrected) for all bimodal conditions. Activations listed with negative contrasts are listed with a negative sign.

Bibliography

- Andersen, T. S., Tiippana, K., and Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Cog. Brain Res.*, 21:301–308.
- Aradi, I. and Soltesz, I. (2002). Modulation of network behaviour by changes in variance in interneuronal properties. *J Physiol*, 538(Pt 1):227–51.
- Arnold, D. H., Johnston, A., and Nishida, S. (2005). Timing sight and sound. *Vis. Res.*, 45:1275–1284.
- Ashburner, J. and Friston, K. (1999). Nonlinear spatial normalization using basis functions. *Hum Brain Mapp*, 7(4):254–66.
- Assisi, C. G., Jirsa, V. K., and Kelso, J. A. S. (2005a). Dynamics of multifrequency coordination using parametric driving: theory and experiment. *Biol. Cybern.*
- Assisi, C. G., Jirsa, V. K., and Kelso, J. A. S. (2005b). Synchrony and clustering in heterogeneous networks wiith global coupling and parameter dispersion. *Phys. Rev. Lett.*, 94:018106.
- Assmus, A., Marshall, J., Ritzl, A., Noth, J., Zilles, K., and Fink, G. (2003). Left inferior parietal cortex integrates time and space during collision judgments. *Neuroimage*, 20 Suppl 1:S82–8.
- Banati, R. B., Goerres, G. W., Tjoa, C., Aggelton, J. P., and Grasby, P. (2000). The functional neuroanatomy of visual-tactile integration in man: A study using positron emission tomography. *Neuropsychologia*, 38:115–124.

- Beek, P. J., Peper, C. E., and Daffertshofer, A. (2002). Modeling rhythmic interlimb coordination: Beyond the haken-kelso-bunz model. *Brain and Cognition*, 48:149–165.
- Benevento, L. A., Fallon, J., Davis, B. J., and Rezak, M. (1977). Auditory-visual interaction in single cells in the cortex of the superior temporal sulcus and the orbital frontal cortex of the macaque monkey. *Exp. Neurol.*, 57:849–872.
- Bertelson, P. and Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Perception and Psychophysics*, 6:578–584.
- Bodner, M., Kroger, J., and Fuster, J. M. (1996). Auditory memory cells in dorsolateral prefrontal cortex. *Neuroreport*, 7:1905–1908.
- Boring, E. (1923). *A history of experimental psychology*. Pendragon, New York.
- Bressler, S. L. (1995). Large scale cortical networks and cognition. *Brain res. rev.*, 20:288–304.
- Bressler, S. L. and Kelso, J. A. S. (2001). Cortical coordination dynamics and cognition. *Trends Cog. Sci.*, 5(1):26–36.
- Buck, J. and Buck, E. (1976). Synchronous fireflies. *Scientific American*, 234:74–85.
- Buhmann, J. (1989). Oscillations and low firing rates in associative memory neural networks. *Phys. Rev. A*, 40(7):4145–4148.
- Bushara, K. O., Graffman, J., and Hallet, M. (2001). Neural correleates of auditory-visual stimulus onset asynchrony. *J. Neurosci.*, 21(1):300–304.
- Byblow, W. D., Carson, R. G., and Goodman, D. (1994). Expressions of asymmetries and anchoring in bimanual coordination. *Human Movement Science*, 13:3–28.
- Calvert, G. A. (2001). Crossmodal processing in the human brain: Insights from functional neuroimaging studies. *Cerebral Cortex*, 11:1110–1123.
- Calvert, G. A., Brammer, M. J., Bullmore, E. T., Campbell, R., Iverson, S. D., and David, A. S. (1999). Response amplification in sensory specific cortices during crossmodal binding. *Neuroreport*, 10:2619–2623.

- Calvert, G. A., Hansen, P. C., Iverson, S. D., and Brammer, M. J. (2001). Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the bold effect. *Neuroimage*, 14:427–438.
- Collins, J. J. and Stewart, I. N. (1993). Coupled nonlinear oscillators and the symmetries of animal gaits. *Journal of Nonlinear Science*, 3:349–392.
- Cox, R. W. (1996). Afni: Software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.*, 29(3):162–173.
- Cross, M. C. and Hohenberg, P. (1993). Pattern formation out of equilibrium. *Revs. of Modern Physics*, 65:851–1112.
- da Silva, F. L. (1991). Neural mechanisms underlying brain waves: from membranes to networks. *Electroencephalogr. Clin. Neurophysiol.*, 79:81–93.
- DeGuzman, G. C., Jirsa, V. K., Assisi, C. G., and Kelso, J. A. S. (2002). Phase and amplitude mediated transitions in coordination phenomena. Society for Neuroscience. 2002 Abstract Viewer/Itinerary Planner.
- DeGuzman, G. C. and Kelso, J. A. S. (1991). Multifrequency behavioral patterns and the phase attractive circle map. *Biological Cybernetics*, 64:485–495.
- Desimone, R. and Cross, C. G. (1979). Visual areas in the temporal cortex of the macaque. *Brain Res.*, 178:363–380.
- Duhamel, J. R., Colby, C. L., and Goldberg, M. E. (1991). *Brain and Space*, chapter Congruent representations of visual and somatosensory space in single neurons of monkey ventral intraparietal cortex (area VIP)., pages 223–226. Oxford University Press.
- Engel, A. K., Fries, P., and Singer, W. (2001). Dynamical predictions: Oscillations and synchrony in top down processing. *Nat. Rev. Neurosci.*, 2:704.
- Ermentrout, B. (1998). Neural networks as spatio-temporal pattern forming systems. *Rep. Prog. Phys.*, 61:353–430.

- Falchier, A., Clavagnier, S., Barone, P., and Kennedy, H. (2002). Anatomical evidence of multimodal integration in the primate striate cortex. *J. Neurosci.*, 22(13):5749–5759.
- Felleman, D. J. and Essen, D. C. V. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1:1–47.
- Fendrich, R. and Corballis, P. M. (2001). The temporal cross-capture of audition and vision. *Perception and Psychophysics*, 63(4):719–725.
- Finger, S. (2001). *Origins of Neuroscience: A history of explorations into brain function*. Oxford University Press.
- Fink, P. W., Foo, P., Jirsa, V. K., and Kelso, J. A. S. (2000). Local and global stabilization of coordination by sensory information. *Experimental Brain Research*, 134:9–20.
- Fitzhugh, R. (1961). Impulses and physiological states in theoretical models of nerve membrane. *Biophys. J.*, 1:445–466.
- Foxe, J. J. and Schroeder, C. E. (2005). The case for feedforward multisensory convergence during early cortical processing. *Neuroreport*, 16(5):419–423.
- Foxe, J. J. and Simpson, G. V. (2002). Flow of activation from v1 to frontal cortex in humans. *Exp. Brain Res.*, 142:139–150.
- Freeman, W. J. (1992). Tutorial on neurobiology: From single neurons to brain chaos. *Int. J. Bif. Chaos*, 2:451–482.
- Friston, K., Holmes, A., Poline, J., Grasby, P., Williams, S., Frackowiak, R., and Turner, R. (1995). Analysis of fmri time-series revisited. *Neuroimage*, 2(1):45–53.
- Friston, K., Holmes, A., and Worsley, K. (1999). How many subjects constitute a study? *Neuroimage*, 10(1):1–5.
- Fuchs, A., Jirsa, V. K., and Kelso, J. A. S. (1999). *Analysis of neurophysiological brain functioning*, chapter Traversing scales of organization 2: Analysis and reconstruction, pages 90–106. Springer Series in Synergetics. Springer-Verlag.

- Fuchs, A., Kelso, J. A. S., and Haken, H. (1992). Phase transitions in the human brain: Spatial mode dynamics. *Int. J. of Bifurcation and Chaos*, 2:917–939.
- Giard, G. H. and Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience*, 11(5):473–490.
- Golomb, D., Hansel, D., Shraiman, B., and Sompolinsky, H. (1992). Clustering in globally coupled phase oscillators. *Phys. Rev. A*, 45(6):3516–3530.
- Grossberg, S., Pribe, C., and Cohen, M. A. (1997). Neural control of interlimb oscillations. *Biological Cybernetics*, 77(2):131–140.
- Haken, H. (1983). *Synergetics: An Introduction*. Springer Series in Synergetics. Springer-Verlag, 3 edition.
- Haken, H., Kelso, J. A. S., and Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological Cybernetics*, 51:347–356.
- Haken, H., Peper, C. E., Beek, P. J., and Daffertshofer, A. (1996). A model for phase transitions in human hand movements during multifrequency tapping. *Physica D*, 90:179–196.
- Hindmarsh, J. and Rose, R. (1994). A model for rebound bursting in mammalian neurons. *Philos Trans R Soc Lond B Biol Sci*, 346(1316):129–50.
- Hubel, D. H. and Wiesel, T. N. (1965). Receptive fields and functional architecture in two non-striate visual areas (18 and 19) of the cat. *J. Neurophysiol.*, 28:229–289.
- Jirsa, V. K., Fink, P., Foo, P., and Kelso, J. A. S. (2000). Parametric stabilization of bimanual coordination: A theoretical model. *Journal of Biological Physics*, 1:85–112.
- Jirsa, V. K., Fuchs, A., and Kelso, J. A. S. (1998). Connecting cortical and behavioral dynamics: Bimanual coordination. *Neural Computation*, 10:2019–2045.
- Jirsa, V. K. and Haken, H. (1997). A derivation of a macroscopic field theory of the brain from quasi-microscopic neural dynamics. *Physica D*, 99:503–526.

- Jirsa, V. K. and Kelso, J. A. S. (2005). The excitator as a minimal model for the coordination dynamics of rhythmic movement generation. *Journal of motor behavior*, 37(1):35–51.
- Jones, E. G. and Powell, T. P. (1970). An anatomical study of converging sensory pathways within the cerebral cortex of the monkey. *Brain*, 93:793–820.
- Jung, T.-P., Makeig, S., McKeown, M. J., Bell, A. J., Lee, T.-W., and Sejnowski, T. J. (2001). Imaging brain dynamics using independent component analysis. *Proc. IEEE*, 89(7):1107–1122.
- Kandel, E. R., Schwartz, J. H., and Jessell, T. M. (1991). *Principles of neural science*. Elsevier.
- Kay, B. A., Kelso, J. A. S., Saltzman, E. L., and Schöner, G. (1987). Steady state and perturbed rhythmical movements: A dynamical analysis. *Journal of Experimental Psychology: Human Perception and Performance*, 17:183–197.
- Kay, B. A. and Warren, W. H., J. (1998). A dynamical model of coupling between posture and gait. In Rosenbaum, D. A. and Collyer, C. A., editors, *Timing of Behavior*. MIT Press.
- Kay, B. A. and Warren, W. H. (2001). Coupling of posture and gait: Mode locking and parametric excitation. *Biological Cybernetics*, 85:89–106.
- Kelso, J., Fuchs, A., Lancaster, R., Holroyd, T., Cheyne, D., and Weinberg, H. (1998). Dynamic cortical activity in the human brain reveals motor equivalence. *Nature*, 392(6678):814–8.
- Kelso, J. A. S. (1984). Phase transitions and critical behavior in human bimanual coordination. *American Journal of Physiology: Regulatory, Integrative and Comparative*, 15:R1000–R1004.
- Kelso, J. A. S. (1995). *Dynamic patterns: The self-organization of brain and behavior*. MIT Press, Cambridge, MA.

- Kelso, J. A. S., Bressler, S. L., Buchanan, S., Deguzman, G. C., Ding, M., Fuchs, A., and Holroyd, T. (1992). A phase-transition in human brain and behavior. *Physics Letters A*, 169(3):134–144.
- Kelso, J. A. S., DeGuzman, G. C., and Holroyd, T. (1991). *Rhythms in physiological systems*, volume 55 of *Springer Series in Synergetics*, chapter Synergetic dynamics of biological coordination with special with special reference to phase attraction and intermittency., pages 195–213. Springer, Berlin.
- Kelso, J. A. S., DelColle, J. D., and Schöner, G. (1990). *Attention and performance XIII*, chapter Action-perception as a pattern formation process, pages 136–169. Erlbaum, Hillsdale, NJ.
- Kelso, J. A. S., Fuchs, A., and Jirsa, V. K. (1999). *Analysis of neurophysiological brain functioning*, chapter Traversing scales of organization 1: Concepts and experiments, pages 73–89. Springer Series in Synergetics. Springer-Verlag.
- Kelso, J. A. S. and Holt, K. G. (1980). Exploring a vibratory systems-analysis of human movement production. *Journal of Neurophysiology*, 45(5):1183–1196.
- Kelso, J. A. S., Holt, K. G., Rubin, P., and Kugler, P. N. (1981). Patterns of human interlimb coordination emerge from the properties of nonlinear limit cycle oscillatory processes: theory and data. *Journal of Motor Behavior*, 13:226–261.
- Kelso, J. A. S., Scholz, J. P., and Schöner, G. (1986). Nonequilibrium phase transitions in coordinated biological motion: Critical fluctuations. *Physics Letters A*, 118:279–284.
- Kling, J. W. and Riggs, L. A., editors (1971). *Experimental Psychology*. Holt, Rinehart and Winston Inc.
- Kugler, P. N. and Turvey, M. T. (1997). *Information, natural law, and the self-assembly of rhythmic movement*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Kuramoto, Y. (1984). *Chemical oscillations, waves and turbulence*. Springer, Berlin.

- Kwong, K., Belliveau, J., Chesler, D., Goldberg, I., Weisskoff, R., Poncelet, B., Kennedy, D., Hoppel, B., Cohen, M., and Turner, R. (1992). Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proc Natl Acad Sci*, 89(12):5675–9.
- Levitin, D. J., MacLean, K., Chu, L. Y., and Jensen, E. R. (2000). The perception of crossmodal simultaneity. *Int. J. Computing and Anticipatory Sys.*, pages 323–329.
- Luria, A. R. (1980). *Higher cortical functions in man*. Basic Books Inc., New York.
- Macaluso, E., Firth, C., and Driver, J. (2000). Selective spatial attention of vision and touch: Unimodal and multimodal mechanisms revealed by pet. *J. Neurophysiol.*, 83:3062–3075.
- Matthews, P. and Strogatz, S. (1990). Phase diagram for the collective behavior of limit-cycle oscillators. *Phys. Rev. Lett.*, 65(14):1701–1704.
- Mayville, J. M., Fuchs, A., Ding, M., Cheyne, D., Deecke, L., and Kelso, J. A. S. (2001). Event-related changes in neuromagnetic activity associated with syncope and synchronization timing tasks. *Human Brain Mapping*, 14:65–80.
- Meredith, M. A. (2002). On the neuronal basis of multisensory convergence: A brief overview. *Cognitive Brain Research*, 14:31–40.
- Mesulam, M.-M. (1998). From sensation to cognition. *Brain*, 121:1013–1052.
- Mesulam, M. M. and Mufson, E. J. (1982). Insula of the old world monkey. iii: Efferent cortical output and comments on function. *J. Comp. Neurol.*, 212:38–52.
- Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., and Foxe, J. J. (2002). Multisensory auditory-visual interactions during early sensory processing in humans: A high-density electrical mapping study. *Cog. Brain Res.*, 14:115–128.
- Monte, S. D., d’Ovidio, F., and Mosekilde, E. (2003). Coherent regimes of globally coupled dynamical systems. *Phys. Rev. Lett.*, 90:054102.
- Morein-Zemir, S., Soto-Faraco, S., and Kingstone, A. (2003). Auditory capture of vision: examining temporal ventroloquism. *Cognitive Brain Research*, 17:154–163.

- Moutoussis, K. and Zeki, S. (1997). A direct demonstration of perceptual asynchrony in vision. *Proceedings of the Royal Society of London B*, 264:393–399.
- Mufson, E. J. and Mesulam, M. M. (1984). Thalamic connections of the insula in the rhesus monkey and comments on the paralimbic connectivity of the medial pulvinar nucleus. *J. Comp. Neurol.*, 227:109–120.
- Nagashino, H. and Kelso, J. A. S. (1992). Phase transitions in oscillatory neural networks. *Science of Artificial Neural Networks, SPIE*, 1710:278–297.
- Nagumo, J., Arimoto, S., and Yoshizawa, S. (1962). An active pulse transmission line simulating nerve axon. *Proc. IRE*, 50:2061–2070.
- Nichols, S. and Wiesenfeld, K. (1992). Ubiquitous neutral stability of splay-phase states. *Phys. Rev. A*, 45(12):8430–8435.
- Nunez, P. L. (1995). *Neocortical dynamics and human EEG rhythms*. Oxford University Press.
- Ogawa, S., Tank, D., Menon, R., Ellermann, J., Kim, S., Merkle, H., and Ugurbil, K. (1992). Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. *Proc Natl Acad Sci*, 89(13):5951–5.
- Ott, E., So, P., Barreto, E., and Antonsen, T. (2002). The onset of synchronization in systems of globally coupled chaotic and periodic oscillators. *Physica D*, 173:29–51.
- Pearson, R. C., Brodal, P., Gatter, K. C., and Powell, T. P. (1982). The organization of connections between the cortex and the claustrum in the monkey. *Brain Res.*, 234:435–441.
- Pecora, L. M. and Carroll, T. L. (1998). Master stability functions for synchronized coupled systems. *Phys. Rev. Lett.*, 80(10):2109–2112.
- Peper, C. E., Beek, P. J., and Van Wieringen, P. C. W. (1995). Frequency induced phase transitions in bimanual tapping. *Biological Cybernetics*, 73(4):301–309.
- Recanzone, G. H. (2003). Auditory influences on temporal rate perception. *J. Neurophysiol.*, 89:1078–1093.

- Rockland, K. S. (2004). Connectional neuroanatomy: The changing scene. *Brain Res.*, 1000:60–63.
- Rockland, K. S. and Ojima, H. (2001). Calcarine area v1 as a multimodal convergence area. In *Soc. Neurosci. abstracts*, volume 27, page 1342.
- Rockland, K. S. and Ojima, H. (2003). Multisensory convergence in calcarine visual areas in macaque monkey. *Int. J. Psychophysiol.*, 50:19–26.
- Rubin, J. and Terman, D. (2000). Analysis of clustered firing patterns in synaptically coupled networks of oscillators. *J Math Biol*, 41(6):513–45.
- Scholz, J., Kelso, J. A. S., and Schöner, G. (1987). Nonequilibrium phase transitions in coordinated biological motion: critical slowing down and switching time. *Physics Letters A*, 123:390–394.
- Schöner, G., Haken, H., and Kelso, J. A. S. (1986). A stochastic theory of phase transitions in human hand movements. *Biological Cybernetics*, 53:247–257.
- Schöner, G., Jiang, W. Y., and Kelso, J. A. S. (1990). A synergetic theory of quadrupedal gaits and gait transitions. *Journal of Theoretical Biology*, 142(3):359–391.
- Schöner, G. and Kelso, J. A. S. (1988a). Dynamic pattern generation in behavioral and neural systems. *Science*, 239:1513–1520.
- Schöner, G. and Kelso, J. A. S. (1988b). A synergetic theory of environmentally-specified and learned patterns of movement coordination: II. component oscillator dynamics. *Biological Cybernetics*, 58:81–89.
- Schöner, G. and Kelso, J. A. S. (1988c). A synergetic theory of environmentally-specified and learned patterns of movement coordination: I. relative phase dynamics. *Biological Cybernetics*, 58:71–80.
- Schroeder, C. E., Lindsley, R. W., Specht, C., Marcovici, A., Smiley, J. F., and Javitt, D. C. (2001). Somatosensory input to auditory association cortex in the macaque monkey. *J. Neurophysiol.*, 85:1322–1327.

- Sekuler, R., Sekuler, A., and Lau, R. (1997). Sound alters visual motion perception. *Nature*, 385(6614):308.
- Seltzer, B. and Pandya, D. N. (1989). Frontal lobe connections of the superior temporal sulcus in the rhesus monkey. *J. Comp. Neurol*, 281:97–113.
- Shams, L., Kamitani, Y., and Shimojo, S. (2000). What you see is what you hear. *Nature*, 408:788.
- Shipley, T. (1964). Auditory flutter-driving of visual flicker. *Science*, 145(3638):1328–1330.
- Spence, C. and Squire, S. (2003). Multisensory integration: Maintaining the perception of synchrony. *Curr. Biol.*, 13:R519–R521.
- Sporns, O., Gally, J. A., Reeke, G. N., and Edelman, G. M. (1989). Reentrant signalling among simulated neuronal groups leads to coherency in their oscillatory activity. *Proc. Natl. Acad. Sci.*, 86(18):7265–7269.
- Stein, B. E. and Meredith, M. A. (1993). *Merging of the senses*. MIT Press.
- Strogatz, S. H. and Mirollo, R. E. (1993). Splay states in coupled josephsons arrays: Analytical prediction of floquet multipliers. *Phys. Rev. E*, 47(1):220–227.
- Sugita, Y. and Suzuki, Y. (2003). Implicit estimation of sound arrival time. *Nature*, 421:911.
- Tallon-Baudry, C., Bertrand, O., Delpuech, C., and Pernier, J. (1996). Stimulus specificity of phase-locked and non-phase-locked 40 hz visual responses in human. *J. Neurosci.*, 16(13):4220–4249.
- Tesche, C., Uusitalo, M., Ilmoniemi, R., Huottilainen, M., Kajola, M., and Salonen, O. (1995). Signal-space projections of meg data characterize both distributed and well-localized neuronal sources. *Electroencephalogr Clin Neurophysiol*, 95(3):189–200.
- Treffner, P. J. and Turvey, M. T. (1996). Symmetry, broken symmetry, and handedness in bimanual coordination dynamics. *Experimental Brain Research*, 107(3):463–478.

- Vaadia, E., Benson, D. A., Hienz, R. D., and Goldstein, M. H. J. (1986). Unit study of monkey frontal cortex: active localization of auditory and visual stimuli. *J. Neurophysiol.*, 56:934–952.
- van Noorden, L. P. A. S. (1975). *Temporal coherence in the perception of tone sequences*. PhD thesis, Eindhoven University of Technology.
- Wilson, H. R. and Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.*, 12:1–24.
- Yamaguchi, Y. and Shimizu, H. (1984). Theory of self-synchronization in the presence of frequency distribution and external noises. *Physica D*, 11(212-226).
- Yamanishi, J., Kawato, M., and Suzuki, S. (1980). Two coupled oscillators as a model for the coordinated finger tapping by both hands. *Biological Cybernetics*, 37:219–225.
- Yuasa, H. and Ito, M. (1990). Coordination of many oscillators and generation of locomotory patterns. *Biological Cybernetics*, 63:177–184.