Background and Significance

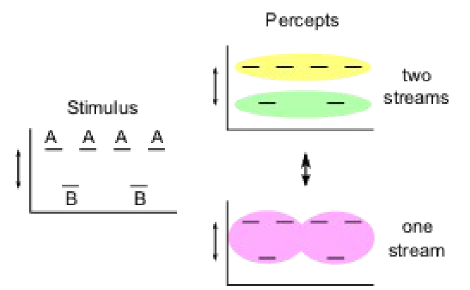


Figure 1: ABA_ stimulus and possible percepts. At the beginning of a trial, subjects hear an integrated percept with a galloping rhythm (bottom). After some time, they will hear a segregated percept (top) and then experience alternations

While researchers have made significant advances in improving hearing sensitivity, the processes that allow listeners to segregate sounds belonging to different sources in a dynamically changing auditory scene are still poorly understood. Our research program, if successful, will both advance basic understanding of how the brain groups and identifies sounds and produce novel techniques that may be useful in experimental analysis and patient assessment.

A critical barrier to progress in the field of auditory scene analysis is the lack of quantitative descriptions of the *dynamic* processes that underlie stream segregation. Current models for the neural basis of stream segregation include grouping by coactivation (Pressnitzer *et al.*, 2008; Pressnitzer *et al.*, 2011), which performs segregation using a discrimination threshold for spike counts, filter-based feedforward models (McCabe and Denham, 1997), and a temporal coherence model (Elhilali *et al.*, 2009). While these methods are useful for determining the steady-state likelihood of segregation for fixed stimuli, they address neither the time-varying changes in segregation probability nor in the stimulus.

Of particular interest is the phenomenon called buildup, for which probability of segregation for an unchanging repeated stimulus increases with time from stimulus onset to a steady state value. The standard qualitative account of buildup cites evidence accumulation as the driving factor in the evolution of segregation probability (Bregman, 1990). We introduce novel applications of dynamical models that can quantitatively describe how buildup occurs. Our statistical model (**Aim 1**) provides an alternative to the accumulation account-- we attribute buildup to alternations between bistable perceptual states. The observation of perceptual bistability (Pressnitzer & Hupe, 2006) for ambiguous auditory stimuli enables the application of the alternating renewal process model, and is also indicative of a competitive neuronal architecture (Wilson and Cowan, 1972; Laing *et al.*, 2010). We are developing and applying ad-hoc neuronal-like models (**Aim 2**) to describe the interactions between neural populations (eventually networks) representing different groupings of the auditory scene. These simulations also reproduce the dynamic changes in likelihood of segregation for unchanging auditory input. The assumption that the neural networks involved undergo bistability allows us to predict how subjects should perceive dynamically changing auditory stimuli, and the effects of previous context (**Aim 3**)-- for instance, because of hysteresis, perception in a dynamically changing scene depends on previous stimuli and perceptual state. These results represent a conceptual overhaul of the buildup process, and advance the field's understanding of the neural bases of stream segregation.

Along with these conceptual advances, the proposed research should ease methodological difficulties in studying buildup specifically and auditory scene analysis in general. The study of neurophysiological mechanisms of stream segregation is difficult because stimuli and timescales are poorly suited for fMRI, with its loud environment and low temporal resolution. ERP/MEG are more appropriate in timescale, but requires averaging over many events and are unsuited for elucidating trial-by-trial effects. Electrophysiology on animals is useful but limited by the ability of animal models to reliably respond on the basis of their higher level perceptual state. By using subjective reports from psychophysical experiments to inform computational models, we can circumvent these difficulties. Our statistical model (**Aim 1**) allows easy interconverting between a number of data types-- from duration distributions to buildup functions, or vice versa. It allows comparison between data from short or long trials, with fixed or changing parameters. Our mechanistic models (**Aim 2**)**,** while presently remote from electrophysiological measurements, will enable better estimates for the strength of cues for segregation and integration, intrinsic noise, and characteristic timescales of processes like adaptation, inhibition, temporal integration and memory. We expect the results of our experiments (**Aim 3**) to yield techniques for estimating dynamic features of higher level auditory processing. In addition, while presently optimized for application to psychophysical data, our techniques are in principle appropriate for any measurement of a buildup function.

Approach/Innovation: Alternating Renewal Process (ARP): The central unifying theory behind the proposed research is that buildup is characterized by an alternating renewal process, as follows:
Suppose that, for an ambiguous ABA_ stimulus, listeners alternate back and forth between hearing integration

and segregation, that the durations of these percepts will be random and independent, and can be described by stationary gamma distributions. Thus the perceptual timecourse for a given trial with a fixed stimulus, eg the reports for being in one or the other of the possible perceptual states over time, can be approximated as alternating draws from each gamma distribution. Finally, if we assume that at the beginning of a trial, the first perceptual state will be that of integration, we can appreciate that buildup is simply the averaging over many trials of random switches out of (as well as back into) a known starting state (Figure 2). This describes an alternating renewal process (ARP).

There are a number of advantages to characterizing the buildup function in this way. First, there exists an analytical solution (Stinchcombe *et al.*, 2012) relating the parameters of the two gamma distributions for percept duration to the buildup function, if the first percept is known, so it is possible to interconvert between buildup functions and the statistics of the dominance durations for each percept. Second, using Monte Carlo simulations, it is possible to generate estimates of the segregation probability for when the initial state is not known; as when the parameters change mid-trial. The four-parameter expression relating probability of segregation over time to distributions of percept durations allows comparison of perceptual dynamics across experimental conditions. And, it challenges the viewpoint that buildup of stream segregation is a gradual accumulative process.

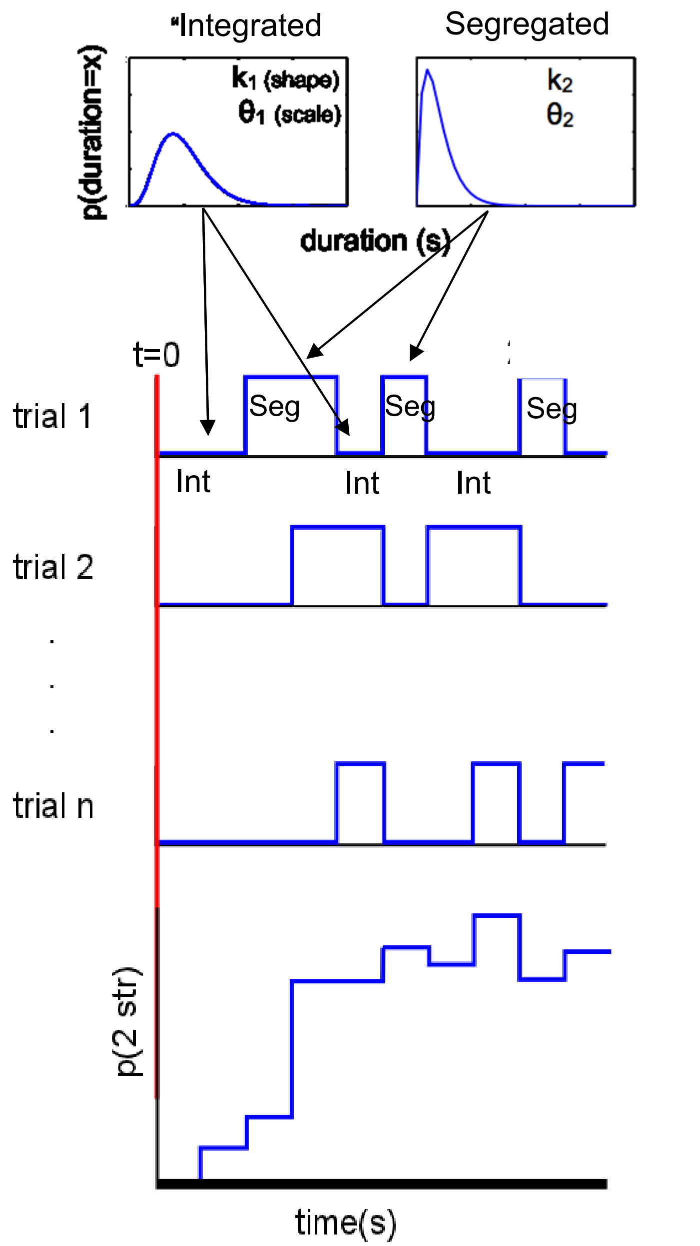


Figure 2: Illustration of alternating renewal process showing how the buildup function (bottom) arises simply by averaging over trials in which the system alternates randomly (according to gamma distributions, top) between two states, with the first state fixed

4. Preliminary Studies

The results of pilot studies in the application of this model to psychophysical data and competition model simulations were presented as a poster at Society for Neuroscience annual conference in 2012 (Steele *et al*, 2012).

Stimuli: We used a repeating ABA_ stimulus with a HLH pattern, such that the A tones were at a higher frequency than the B tones. The tones and the silent interval after each triplet were 125 ms duration, and there were no gaps between adjacent tones. The frequency difference between the tones, DF, was varied between 3 and 7 semitones (st).

Study 1- Long Presentations: To obtain an estimate of the probability distributions of the durations for both integrated and segregated percepts, as well as a characterization of steady state dynamics, repeating ABA_ sequences were presented for 4 minutes (480 triplets). Subjects were instructed to press and hold one key when they heard an integrated percept, and to hold down another key when they heard a segregated percept. Stimulus presentation and response collection were implemented in MATLAB. Two subjects completed three trials each for multiple DF conditions.

Analysis: The times of the keyholds for integrated and segregated percepts were used to obtain percept durations and experimental timecourses. Each experimental timecourse consisted of zeros and ones with 1 ms bins reflecting whether the subject was reporting integration or segregation, respectively. The percept being reported when the trial ended was not included in analysis, as it would have been artificially truncated. We obtained estimates of the parameters for the gamma distributions for integrated and segregated percept durations using maximum likelihood estimation (MLE).

To obtain an estimate of the buildup function with only three repetitions of each stimulus condition, we used a novel method we have dubbed “switch-triggered averaging” (Figure 3). Since we know that the first percept for any given trial is (nearly) always going to be that of integration, we can construct an estimate for the buildup function by averaging over switches into integration, instead of averaging over trials. This is effectively the same as treating every switch into the integrated percept as the beginning of a new trial, and allows us to estimate what the buildup function would look like for many trials (even though we only have three for each subject/condition).

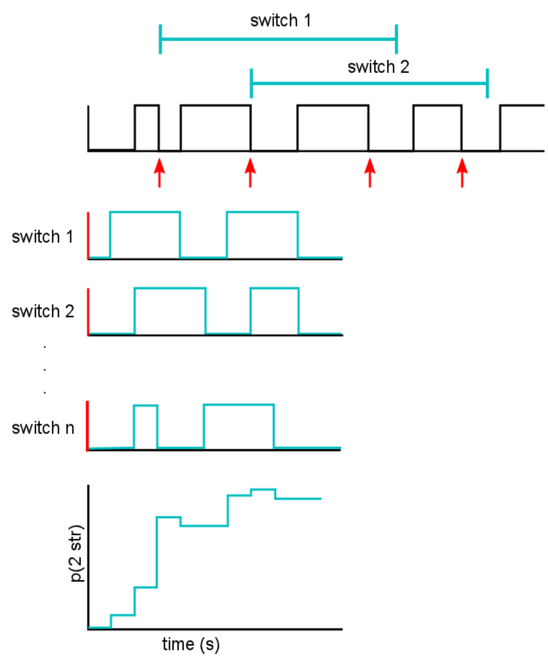


Figure 3: Illustration of switch triggered averaging method for constructing buildup functions from long trials. Averaging aligned by switch into the integrated percept.

<u>Results:</u> The ARP model buildup function using the MLE obtained gamma parameters provided excellent fits for the switch-triggered average obtained buildup function, so long as the listener reported a reasonable number of samples of the duration of each percept, ie, the stimulus was sufficiently ambiguous (Figure 4). Note that the ARP prediction captures not only the asymptotic likelihood of segregation, but that it also does a remarkable job of capturing the dynamics of the change in segregation likelihood over time from the beginning of integration percept epochs.

<u>Study 2- Short presentations</u>: We wanted to establish whether buildup functions constructed in the typical way, by averaging over trials, could be well characterized by the ARP model. To ensure that we could capture a reasonable number of alternations and full durations for both integrated and segregated percepts, we presented a strongly ambiguous repeating ABA_stimulus with DF = 5. Each trial was 20 s long, and there were 100 trials. One subject participated in this experiment (not the author or the sponsor).

<u>Results:</u> The buildup function obtained by both averaging across trials and by using the MLE gamma parameters for the percept duration distributions for the ARP model were very similar (Figure 5). Thus it appears that our ARP model is able to capture the dynamics of stream segregation from the beginning of the onset of an ambiguous sound sequence.

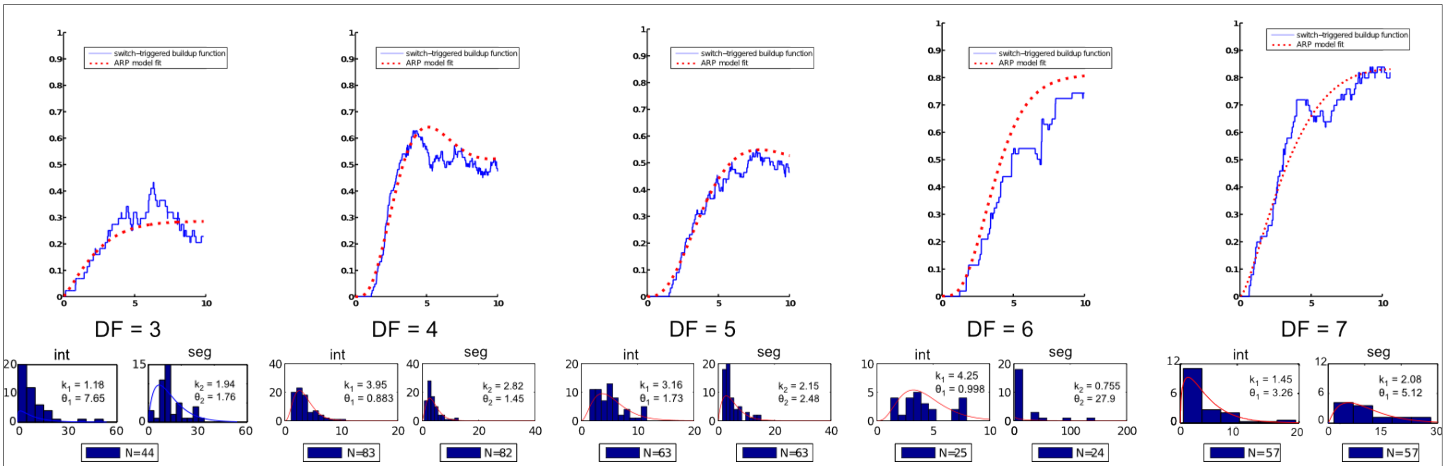


Figure 4: Results from long presentations. The blue curves show the buildup function obtained through switch-triggered averaging, while the red curves are obtained from the ARP model and maximum likelihood parameters for the duration distributions for integrated and segregated percepts (bottom, histograms). One subject was tested on DF = 3 and DF = 7, and the other conditions were tested in another subject.

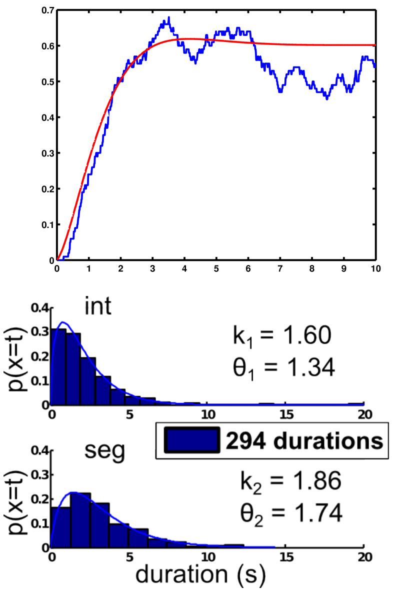


Figure 5: Results from short trials. Buildup was computed by averaging across trials (top, blue curve) as well as by analytical solution to the ARP model using MLE gamma parameter fits for duration distributions (bottom)

Study 3. Competition model

Previous work from our lab has produced a competition model (Shpiro et al, 2009) that produces gamma-distributed dominance durations. This model was originally created to recreate the dynamics of visual bistability, in particular, perception of ambiguous motion displays, or plaids. Plaid stimuli are made from two drifting gratings, and under ambiguous conditions will also produce alternating percepts between coherence (analogous to integration, in which the gratings appear to fuse and travel in one direction) and transparency (analogous to segregation, in which each grating's motion direction is perceived separately).

This model uses two neural populations, one which represents a "segregated" sound organization and the other representing "integrated" percept. These two populations mutually inhibit each other (Figure 6). A tacit assumption with this model is that perception corresponds to a high firing rate in one of these populations, and perceptual dominance occurs when one population's firing rate is higher than that of the other-- because of mutual inhibition, the dominant population will suppress the other. Alternations between perceptual dominance states under this model can be driven by both noise and adaptation. We conducted simulations under a number of stimulation regimes, including matched vs mismatched input currents to each population. We simulated 5000 s or enough time to collect roughly 1000 durations.

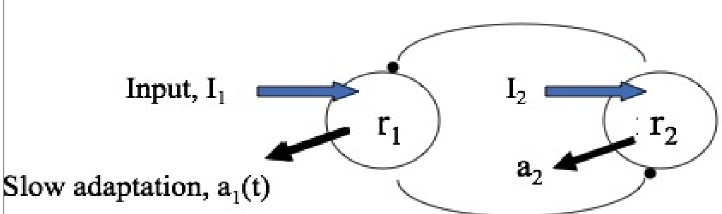
Analysis: The simulated timecourses of population firing rates were converted to binary timecourses relating when one population's firing rate was higher than the other's, and from these we calculated the distribution of dominance durations. The population with the first dominance epoch was arbitrarily labelled "coherent," except when the input current to each population was mismatched. In this case, one population was always labelled "coherent", and simulation runs in which the other population attained dominance first by chance were discarded. We averaged over switches into integration to produce buildup functions, and used MLE gamma parameters from the dominance duration distributions to generate ARP predictions.

Results: For all simulations that produced sufficient alternations, the ARP model provided a good description of the buildup function obtained from competition model simulations. This is especially impressive because in adaptation-driven alternation regimes in the competition model, first-order correlations are observed in percept-to-percept durations. The presence of dependence between the durations of successive percepts violates the assumptions of our renewal process, but the ARP model generates good descriptions anyway (Figure 6).

5. Research design and methods

Specific Aim 1: An alternating renewal process (ARP) can account for the time course of *buildup*, the increased probability of stream segregation over time. The model suffices without invoking any mechanism for evidence accumulation.

Overview: While preliminary results from studies 1 and 2 show that the gamma parameters for percept duration distributions, through the ARP model, strongly determine the buildup function obtained empirically, we wish to



$$\tau\, dr_1/dt = -r_1 + S(-\beta r_2 - \varphi\, a_1 + I_1)$$
$$\tau_a\, da_1/dt = -a_1 + f_a(r_1)$$

$$\tau\, dr_2/dt = -r_2 + S(-\beta r_1 - \varphi\, a_2 + I_2)$$
$$\tau_a\, da_2/dt = -a_2 + f_a(r_2)$$

$$\tau_a >> \tau, \quad S(u)=1/(1+\exp[(\theta-u)/k])$$

Figure 6: Mutual inhibition model framework (Hugh R Wilson, 2003).
Two populations mutually inhibit each other with connection strength Beta. Input currents I1 and I2 are supplied directly to each population, which adapt (a) at a slow timescale (tau). Phi represents gain on adaptation. Theta and k define threshold and slope of input-output function. For specific parameters see Shpiro et al, 2009.

test its analytical power, and compare different experimental conditions in the same subjects.

Subjects: Ten healthy adults will participate in the experiments. They will be screened for normal hearing thresholds and family history of hearing loss.

Stimuli: We will use the same stimuli as in preliminary studies 1 & 2 (Figure 1). We will use both short (20 s) and long (4 min) presentations.

Behavioral task and analysis: Subjects will hold down one button for “integration” and another button for “segregation”. There will be 50 trials for short presentations, and 3 trials for long presentations, collected over two sessions with frequent break intervals.

We will collect responses first for short presentation, and average over trials and subjects to construct buildup curves for each DF condition. We will investigate the feasibility of applying the model in the reverse direction from our preliminary studies, searching for gamma parameters that minimize least squares error between the ARP prediction and the empirically obtained trial averaged buildup function. We will use the fits to the buildup function obtained from short trials to predict the dynamic behavior for long trials, and then test with long trials (4 minutes, 3 trials).

Subjects' reports will be converted into binary timecourses with 1 ms bins, and averaged over trials (short presentations) or over switches into integration (long presentations) to produce buildup functions. In particular, **fitting directly to the buildup function obtained by averaging over short trials should yield parameters for the gamma distributions reflecting percept durations in long trials**. Because we will have both forms of data, we will be able to directly compare the results obtained by fitting from the buildup function to derive gamma distributions as well as what we have previously demonstrated, the description obtained by fitting the gamma parameters and estimating the buildup function. Such a result would greatly extend to applicability of the ARP model, as it could be used on any buildup function to obtain estimates of the underlying percept duration distributions.

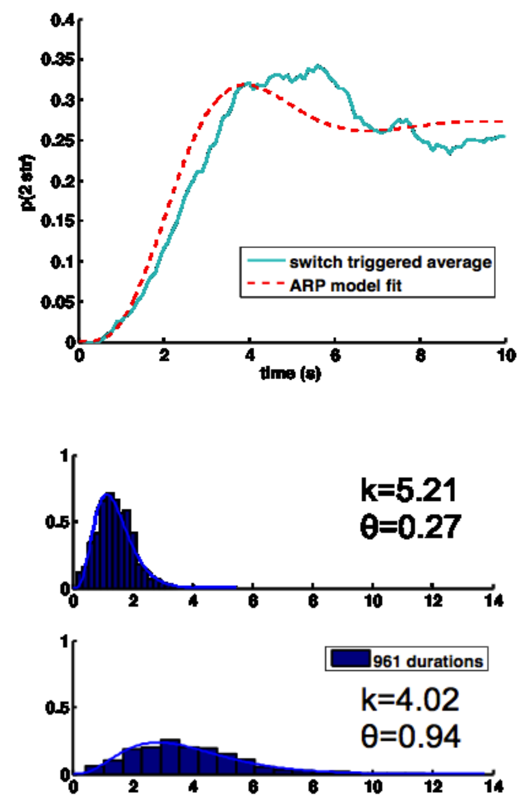


Figure 7: Competition model produces a plausible buildup function for a stimulus that biases the listener towards integration, or hearing galloping triplets.

Potential Pitfalls: We had early concerns with the application of this approach to data acquired from human listeners using ambiguous ABA_ stimuli, in particular because the assumption of stationarity might fail. A number of studies have shown that the duration of the first percept is typically longer than subsequent percept It would be relatively trivial and potentially illuminate the ARP with three distributions using Monte Carlo. It would be relatively trivial to introduce a third gamma distribution to this model, namely that describing the distribution of durations for the first integrated percept, to produce monte carlo estimates of the buildup functions. What is compelling is the degree to which the model works even without this consideration.

Specific Aim 2: Neural competition networks are sufficient to reproduce psychophysical results as well as the behavior of the ARP model.

Overview: Preliminary results from study 3 show that existing two-population competition models can capture the dynamics of buildup, even when assumptions such as independence between dominance durations are violated. This model parameterizes the connection weight of mutual inhibition between the two populations; strength, form, and timescale of adaptation within each neural population; the input driving each population; the contribution and timescale of noise; and the rate of growth and decay of neural activity. After I learn the techniques of dynamical systems analysis from my sponsor, I will be able to modify these models to pursue the following goals:

SA 2.1: Find mappings of parameters for mechanistic models onto specific gamma distributions and buildup functions obtained experimentally. Since we have already verified that our first pass rendition of the reduced neuronal-like model is consistent with both behavior and our new ARP model, the next logical step is to try to unify all three of these. We will conduct basic parameter searches to determine the relationship between the four parameters that describe buildup behavior and alternation statistics and the free parameters of the competition models. The expected result is that we will be able to match mechanistic parameters that reliably reproduce behavioral results, suggesting potential neurobiological mechanisms for the processes underlying stream segregation.

SA 2.2: Define inputs to each population in the competition model by computations performed on stimulus histogram. One thing which we would like is to know how alternation dynamics depend on the stimulus characteristics, in addition to internal processes of competition dynamics. To evaluate how the stimulus drives each population, we intend to add an input layer to the model. This input layer will take in a spectrogram of the stimulus over time and calculate the strength of the cues for integration and the strength of the cues for segregation. These calculated values will be used for the input current for each separate population.

SA 2.3: Enhance models based on results of behavioral experiments: We expect that our first pass mechanistic models may not be able to reproduce all effects we see behaviorally, such as maintenance of perceptual state during silence (**Aim 3**). Thus, we will update our model with appropriate mechanisms, such as recurrent excitation (like in short term memory models), to synergistically relate to the empirical results we find.

Specific Aim 3: The effects of previous exposure and context will lead to significant changes in both buildup and alternation statistics. The models will be applied and enhanced in order to account for context effects.

Materials and methods: We will use the same subjects as in Aim 1. These experiments will be conducted using the same kind of ABA_ stimuli. Dynamic stimuli will vary DF continuously throughout a trial.

Experiment 1: Dynamic stimuli will be 120 second sequences of repeating ABA_ with DF slowly varying over the range from 3 to 7 st, the same as those used in previous experiments. These stimuli will be arranged so that DF is either ascending to descending, or descending to ascending. There will be 30 trials for each sequence.

Analysis: We will generate predictions for likelihood of segregation as a function of time in sequence, and evolving dF values, using the gamma parameters we obtain in Experiment 1. We will use Monte Carlo simulations based on random samples from these and interpolated distributions (for DF values between those previously tested), using the ARP model to generate percept duration samples composing simulated timecourses. These simulated timecourses will be averaged and compared to the averaged experimental timecourses to determine whether probability of stream segregation in a dynamic scene can be predicted by the steady state statistics of fixed scenes.

Experiment 2: The DF will be chosen based on the results of experiment 1, such that both integrated and segregated percepts are possible and buildup is not too fast or slow, but likelihood of segregation at the end of the trial is high, eg 7 semitones. When a trial ends, one of five silent inter-trial intervals (ITI) will intervene-- 1.2, 2.4, 3.6, 4.8, and 6 seconds. We expect that the “reset” of buildup will strongly depend on the length of the silent interval between trials.

Experimental timecourses will be averaged to construct buildup functions for each preceding ITI. We will then calculate the change in probability of segregation during the first 1 s as a function of ITI, and use that function to fit parameters for mechanistic model simulations.

Potential Pitfalls: We need to ensure that we vary the stimulus at an appropriate timescale-- if the stimulus varies faster than the intrinsic dynamics of perceptual alternations, we will likely lose predictive power by simple Monte Carlo simulation. However, we may be able to propose a rule for switching out of a perceptual epoch when the underlying parameters change.