



# Stimulus Pauses and Perturbations Differentially Delay or Promote the Segregation of Auditory Objects: Psychoacoustics and Modeling

James Rankin<sup>1,2\*</sup>, Pamela J. Osborn Popp<sup>2</sup> and John Rinzel<sup>2,3</sup>

<sup>1</sup> Department of Mathematics, University of Exeter, Exeter, UK, <sup>2</sup> Center for Neural Science, New York University, New York, NY, USA, <sup>3</sup> Courant Institute of Mathematical Sciences, New York, NY, USA

## OPEN ACCESS

### Edited by:

Robert J. Zatorre,  
McGill University, Canada

### Reviewed by:

Yi Du,  
Institute of Psychology (CAS), China  
Claude Alain,  
Rotman Research Institute, Canada

### \*Correspondence:

James Rankin  
james.rankin@gmail.com

### Specialty section:

This article was submitted to  
Auditory Cognitive Neuroscience,  
a section of the journal  
Frontiers in Neuroscience

Received: 14 November 2016

Accepted: 23 March 2017

Published: 20 April 2017

### Citation:

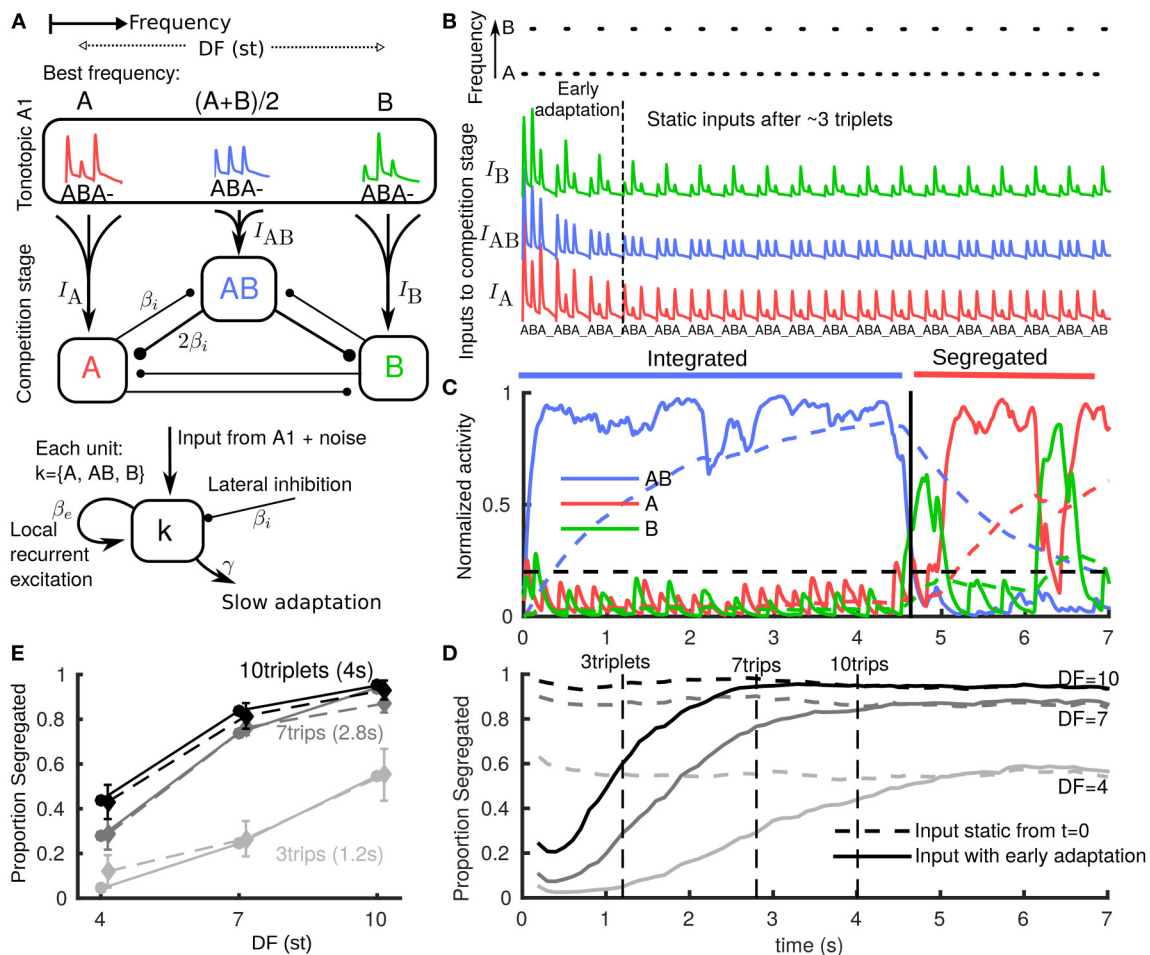
Rankin J, Osborn Popp PJ and  
Rinzel J (2017) Stimulus Pauses and  
Perturbations Differentially Delay or  
Promote the Segregation of Auditory  
Objects: Psychoacoustics and  
Modeling. *Front. Neurosci.* 11:198.  
doi: 10.3389/fnins.2017.00198

Segregating distinct sound sources is fundamental for auditory perception, as in the cocktail party problem. In a process called the build-up of stream segregation, distinct sound sources that are perceptually integrated initially can be segregated into separate streams after several seconds. Previous research concluded that abrupt changes in the incoming sounds during build-up—for example, a step change in location, loudness or timing—reset the percept to integrated. Following this reset, the multisecond build-up process begins again. Neurophysiological recordings in auditory cortex (A1) show fast (subsecond) adaptation, but unified mechanistic explanations for the bias toward integration, multisecond build-up and resets remain elusive. Combining psychoacoustics and modeling, we show that initial unadapted A1 responses bias integration, that the slowness of build-up arises naturally from competition downstream, and that recovery of adaptation can explain resets. An early bias toward integrated perceptual interpretations arising from primary cortical stages that encode low-level features and feed into competition downstream could also explain similar phenomena in vision. Further, we report a previously overlooked class of perturbations that promote segregation rather than integration. Our results challenge current understanding for perturbation effects on the emergence of sound source segregation, leading to a new hypothesis for differential processing downstream of A1. Transient perturbations can momentarily redirect A1 responses as input to downstream competition units that favor segregation.

**Keywords:** auditory perception, psychoacoustics, auditory stream segregation, dynamical systems, computational neuroscience

## 1. INTRODUCTION

A valued paradigm for studying auditory streaming involves segregating two interleaved sequences of A tones and B tones, distinguishable by a perceived difference in pure tone frequency and timing. The tones are organized in a repeating ABA\_ABA\_... pattern (van Noorden, 1975) (“\_” represents silence) (Figure 1B, top). At first heard as a one stream rhythm (integrated percept), the probability of hearing two streams (segregated percept) gradually builds up over several to tens of seconds (build-up) (Anstis and Saida, 1985; Bregman, 1994; Carlyon et al., 2001). Build-up occurs more rapidly with a large difference in frequency (DF) between A and B and at faster presentation rates.



**FIGURE 1 | Neuromechanistic model captures initial bias to integration and build-up of stream segregation. (A)** Model schematic with two stages: tonotopic A1 and a competition stage (downstream of and pooling inputs from A1). A1 encodes only stimulus features, while the downstream competition stage encodes percepts. Inputs from lower frequency A and higher B tones generate onset-plateau responses in A1 dependent on difference in frequency (DF) in semitones (st). In the competition stage three units encode the integrated percept (AB), the segregated A stream, and the segregated B stream. Units are in competition through mutual inhibition, pool inputs from A1, have recurrent NMDA excitation (timescale 70 ms) and undergo slow adaptation (timescale 1.4 s). **(B)** (top) Stimulus paradigm where low A tones, high B tones and silences (—) each of 100 ms repeat in an ABA<sub>n</sub> triplet pattern. (below) A1 responses to tones adapt rapidly (timescale 500 ms) with tonotopic dependence emerging and overall amplitude reducing during first 2–3 triplets. Vertical offset for visualization only. **(C)** One model simulation showing the activation threshold (horizontal dashed), and each population's NMDA variable (solid, pulsatile inputs appear smoothed in sub-threshold activity) and adaptation variable (dashed). When the central AB unit is active (integrated), the peripheral units are suppressed through mutual inhibition. Increasing adaptation for AB increases the probability of noise inducing a switch; when units A or B become active and dominant after ~4.5 s (segregated), the integrated (AB) unit is suppressed. **(D)** Build-up function computed as time-binned trial-averaged proportion segregated computed from  $N = 500$  model simulations. With no early adaptation of inputs from A1 (input static), there is no build-up and stable proportion segregation from long-term alternations is reflected at onset. Early adaptation of inputs from A1 gives initial bias toward integrated and proportion segregated gradually builds up to DF-dependent value of long-term alternations. **(E)** Snapshots from build-up after 3, 7, and 10 triplets from model (each solid curve in E corresponds to a dashed vertical line in (D)) are compared with psychoacoustic data ( $N = 8$  normal hearing subjects) with percept reported at the end of presentation (dashed curves; errorbars show s.e.m., same for all plots).

However, abrupt change in the incoming sound (e.g., a step change in location, loudness or timing) can reset perception to integrated (Anstis and Saida, 1985; Rogers and Bregman, 1993, 1998), after which multisecond build-up begins again. The first perceptual switch, typically from integrated to segregated, is followed by persistent alternations between the two interpretations (Pressnitzer and Hupé, 2006; Denham et al., 2013). Build-up progresses not just to the segregation, but to the stable probability of segregation in the subsequent long-term alternations.

Neural responses to triplet stimuli have been studied in primary auditory cortex (A1) of awake monkeys (Fishman et al., 2001, 2004; Micheyl et al., 2005), in forebrain of awake (Bee and Klump, 2004; Bee et al., 2010) or behaving (Itani and Klump, 2014) songbirds, and in the auditory periphery of anesthetised guinea pigs (Pressnitzer et al., 2008). The tonotopic organization of A1 and increased forward masking at higher presentation rates (Fishman et al., 2001, 2004) can explain the feature dependence of these responses. Studies comparing neural response data with build-up functions from

human psychoacoustic experiments have shown that a trial averaged neurometric function can be tuned to match trial averaged behavioral data (Micheyl et al., 2005; Pressnitzer et al., 2008; Scholes et al., 2015). However, no study has claimed that the neural substrate for the perceptual state or switches in perceptual states lies in or before A1. Indeed, the only animal study with neural data recorded from behaving animals (Itatani and Klump, 2014) concluded that only stimulus features and not perceptual choice is encoded in songbird forebrain (analogous to A1). Responses to tones in A1 show rapid adaptation in the first few hundred milliseconds (1–3 triplets) (Micheyl et al., 2005). In this initial phase, response amplitude adapts and dependence on DF emerges (at first little tonotopic dependence is evident for tones separated by less than an octave). The relationship between this rapid adaptation (~500 ms) and the slower build-up process (several seconds) remains unexplained.

In Rankin et al. (2015) we developed a neuromechanistic model of auditory bistability based on a conceptual model proposed in Fishman et al. (2001). Far apart A and B tones drive tonotopically segregated representations, but for smaller DF the receptive fields overlap, leading to a common drive for an intermediate population encoding integration (Figure 1A). Our model mimics the periodic, pulsatile responses and stimulus feature dependence from A1 (Micheyl et al., 2005), which are pooled as inputs to a competition stage residing downstream (A1 encodes only stimulus features, not the percepts). At the competition stage peripheral units A and B encode segregation and a central unit AB encodes integrated. The competition network incorporates the mechanisms of mutual inhibition, slow adaptation and additive noise shown to play an important role in perceptual bistability (Moreno-Bote et al., 2007; Shpiro et al., 2009). Recurrent excitation with an NMDA-like timescale links responses and thereby percepts across silent gaps between tones and triplets (Figure 1C). Our model captures the complex dynamics of perceptual alternations, reproducing characteristic features such as the log-normal distribution of perceptual durations as well as dependence of perceptual durations on parameters such as DF (Rankin et al., 2015). We focused previously on the alternations after the first perceptual switch; the initial bias to integrated and build-up were not addressed.

Here, we propose that the initial integration bias is determined by early broad tonotopic tuning of neuronal responses in A1, while the multisecond timescale of build-up is due to slow adaptation downstream of A1. Recovery of early adaptation, say after a stimulus pause, can further explain the reset to the integrated percept. Furthermore, we find in new experiments a class of transient perturbations (single unexpected tones in the ongoing stimulus) that subsequently promote segregation, in contrast to the widely reported resets to integrated. Our model, motivated from neurophysiological studies, provides a mechanistic explanation for build-up and resetting whilst also accounting for new experimental findings.

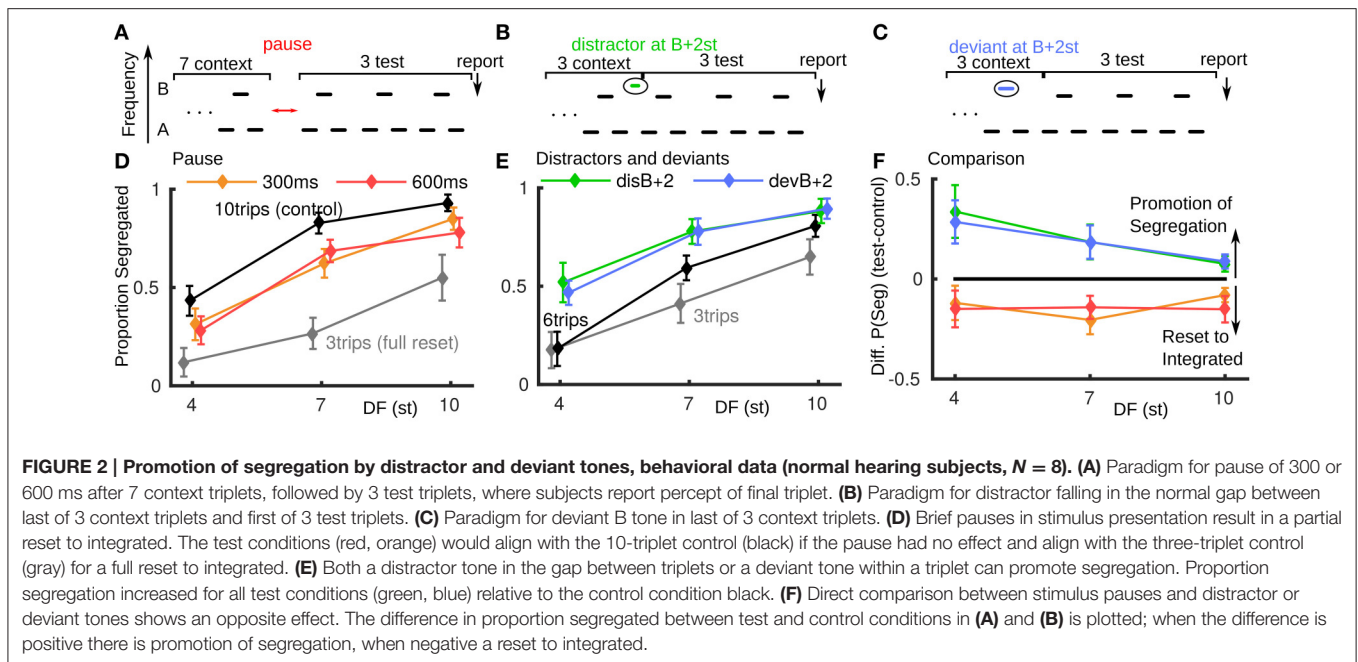
## 2. RESULTS

### 2.1. Neuromechanistic Model Explains Initial Bias to Integration and Build-Up of Stream Segregation

In order to study build-up in our existing model, we made one change to the inputs based on further observations about the early responses to triplets in A1 (Micheyl et al., 2005). We introduced rapid adaptation (timescale 500 ms) for both input amplitude and DF dependence (Figure 1B). During the first 2–3 triplets input evolves as if driven by a DF that is effectively small but gradually increasing to a static value. The AB unit receives enough input bias to become active, suppress the peripheral units and become dominant first (Figure 1C). Time-binned build-up functions (three DF and two input cases) were computed by averaging across simulations. In the input static case (Figure 1D dashed) the inputs are assumed post fast-adaptation (Figure 1B after 3 triplets) and the time-course only reflects the static probability of post build-up alternations. In the input adapting case (Figure 1D solid) responses are initially biased to integrated and gradually build-up to the static probability of later alternations. The slower timescale of this build-up arises from the mechanisms already established in (Rankin et al., 2015) for the competition stage downstream of A1. In particular, there is a slower adaptation process at the competition stage. In psychoacoustic experiments, the build-up process can be sampled with short stimulus presentations of different lengths with percepts reported at the end. Vertical lines in Figure 1D show three such snapshots from the model (Figure 1E solid). These are compared with psychoacoustic data (Figure 1E dashed) for three DF and two presentation length conditions. A repeated measures ANOVA showed a significant effect of DF [ $F_{(2,14)} = 37.49, P < 0.001$ ], of presentation length [ $F_{(2,14)} = 19.49, P < 0.005$ ] and their interaction [ $F_{(4,28)} = 4.34, P < 0.05$ ], see Appendix A in Supplementary Material. The close match with these data show that the model accurately captures build-up (increasing segregation with both DF and presentation length). Our model is the first to produce the bias to integrated in a manner directly motivated from neurophysiology data (Fishman et al., 2001; Micheyl et al., 2005) (fast adaptation in A1) and to produce gradual build-up due to a slower adaptation timescale downstream of A1 (at the competition stage in our model).

### 2.2. Promotion of Segregation by Distractor and Deviant Tones

In psychoacoustic experiments we reproduced a previously reported reset toward integration for a brief pause between triplets (paradigm, Figure 2A; data Figure 2D). In all experiments described here, the stimulus ends in three normal triplets with the last triplet reported as integrated or segregated (Carlyon et al., 2003; Haywood and Roberts, 2010). In Figure 2D, if the test conditions (300 or 600 ms pause) showed no effect, the orange and red curves would align with the black ten triplet control. For a full reset to integrated the



test conditions would align with the three triplet gray curve. Our results show, consistent with existing studies (Beauvois and Meddis, 1997; Denham et al., 2010), that brief stimulus pauses can result in a partial reset back toward integrated. The pause conditions had a significant effect on proportion segregated [ $F_{(2, 14)} = 5.126, P < 0.05$ ], see Appendix A in Supplementary Material. The reset is of a similar magnitude for all pause duration and DF conditions.

In a new experiment six triplet presentations are used with a perturbation in the third triplet (full details in Section 5 and Appendix A in Supplementary Material). In the *distractor case* (Figure 2B), an additional tone is inserted in the normal gap between the third and fourth triplet: ...ABA\_ABAdABA\_..., where “d” is 2 semitones (st) higher than B. In the *deviant case* (Figure 2C), the B-tone in the third triplet is a deviant: ...ABA\_ADA\_ABA\_..., where “D” is 2 st higher than B. A shorter presentation length was used relative to the pause experiment to avoid ceiling effects (saturation at proportion segregated = 1). See Figure 2E, where again, for no effect the test conditions would align with the black control case and for a reset to integrated, move down toward the gray three-triplet case. We found an opposite effect from pauses for a deviant or distractor tone during the ongoing triplet sequence: promotion of segregation. The increase in proportion segregated is significant for these test conditions [ $F_{(3, 21)} = 5.80, P < 0.05$ ]. There is a similar effect for the deviant and distractor cases (largest for small DF). A distractor at 15 st above B showed no effect (not shown); see Appendix A in Supplementary Material.

For each experiment, by calculating the difference in proportion segregated between the test cases (colored curves) and control cases (black curves) in Figures 2D,E, we can make a direct comparison between the two types of perturbation (Figure 2F). A negative (positive) difference indicates a reset

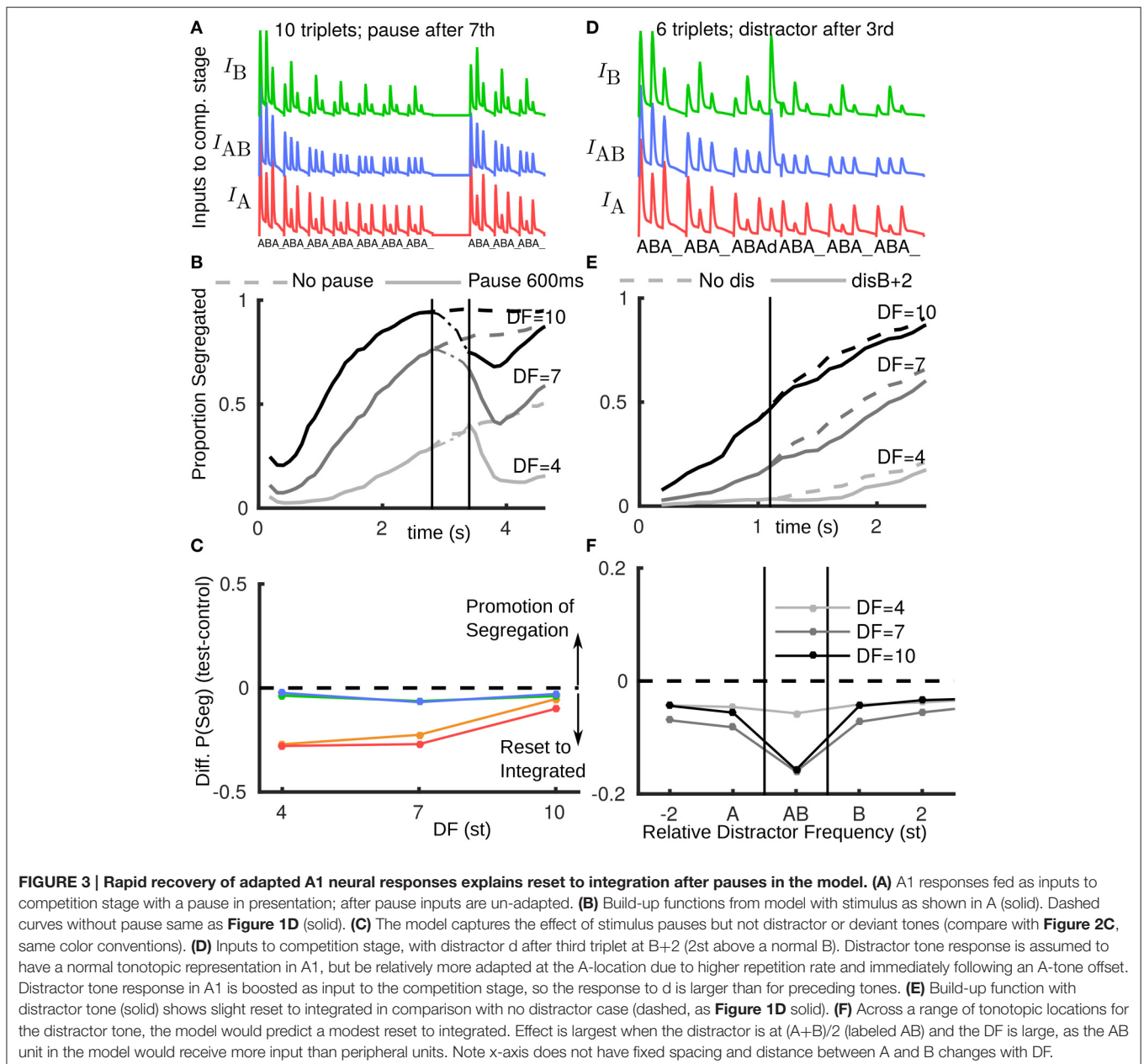
toward integrated (promotion of segregation). The promotion of segregation by a single-tone perturbation during triplets is a new and unexpected finding, opposite to the effect of pauses and other perturbations previously reported. To better understand this phenomenon, we focused on the distractor tones and further investigated their relative frequency to the triplet tones (Figure 4F). Before reporting these data we explore perturbations with the model.

### 2.3. Rapid Recovery of Adapted A1 Responses Explains Reset to Integration for Pauses

In the model we assume that when the stimulus resumes after even a brief pause, it will be partially recovered from adaptation (to a state similar to stimulus onset) (Figure 3A). Figure 3B shows a simulation-averaged build-up function comparing a case without a stimulus pause (input Figure 1B) to a case with a pause input (input Figure 3A). When the stimulus turns off the proportion segregated decreases (increases for DF = 4) toward 0.5. When the stimulus resumes the amplitude and effective DF of inputs from A1 have partially recovered; the proportion segregated continues to decrease (starts decreasing for DF = 4) before resuming gradual build-up. In this way, the model accounts for the partial reset toward integration across all DF conditions, compare red/orange curves in Figure 3E(model) with Figure 2F (experiments).

### 2.4. Model Hypothesis on Differential Processing of Non-triplet Tones

For a distractor tone in the model, in order to compute an input amplitude, we first assumed the same rules as for the standard A and B tones. One modification was to assume a



reduced response in A1 at the A-location due to higher repetition rate and the distractor immediately following an A-tone offset (stimulus-specific adaptation Ulanovsky et al., 2004; Taaseh et al., 2011). Until now the responses in A1 were taken directly as inputs to the competition stage, without modification. However, in initial simulations we found almost no effect of introducing a single new tone. A further assumption is that a distractor tone, arriving in a window where silence was expected, would be detected as a new event, and boosted (approximately to the level of an un-adapted tone) as input to the competition stage. **Figure 3D** shows inputs for a distractor 2 st above a normal B ( $B+2$ ), see Appendix B in Supplementary Material. Still, only a small reset toward integrated is observed (**Figure 3E**). Using the same assumptions for a deviant B tone at  $B+2$  we find a

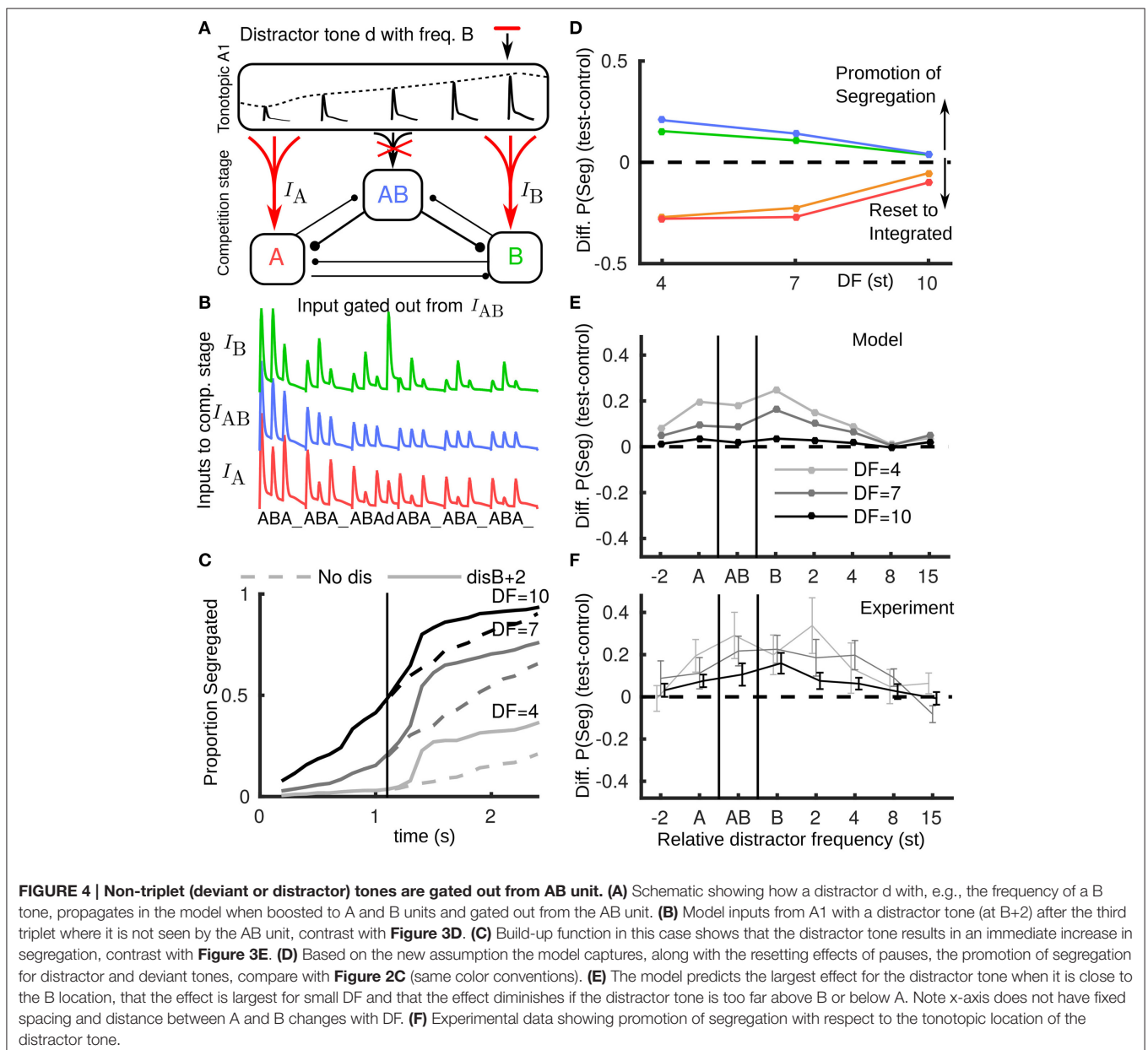
similar effect (**Figure 3C**). A comparison with the experimental data from **Figure 2F** shows that the model has not captured the effects of deviants and distractors. A further exploration varying relative frequency of the distractor tone (**Figure 3F**) shows that the model would predict a large reset toward integrated when it is at a frequency  $(A+B)/2$ , in which case the AB unit receives the most additional input from the distractor tone. However, this prediction was not borne out in later experiments.

Using the model, we tested a new hypothesis for how novel inputs, tones that are saliently not part of a triplet, propagate from A1 to the competition stage. These include tones not fitting the temporal pattern of a regular triplet (e.g., the distractor tone) or not matching the frequency of the tones in a regular triplet (e.g., a deviant tone); in informal listening either case is

saliently different from a normal triplet. We suppose that the AB unit, encoding the integrated percept, will only receive inputs matching a normal triplet, while as before, the unexpected event results in boosted input to the competition stage (Figure 4A). For example, a distractor tone B+2 leads to a larger than expected input at B, but no input to AB (Figure 4B). The build-up function shows an increase in segregation (Figure 4C) due to the peripheral units receiving more input. In both the distractor and deviant cases segregation is promoted, recapitulating the behavior with the reported experimental data, compare Figure 4D (model) with Figure 2F (experiments). Note that the model captures the largest promotion of segregation occurring for small DF.

We further applied the model to predict the dependence of change in proportion segregated on the frequency of a distractor

tone (Figure 4E). Predictions: (1) the promotion of segregation occurs for a range of relative frequencies for the distractor tone, (2) the effect is strongest when the distractor tone is close to the A and B tones, (3) there is no effect if the distractor is too far in tonotopy from the A and B tones, and (4) asymmetry, e.g., that the effect is more prominent when the distractor is near or above the B tone than when it is near or below the A tone. Further experiments confirmed these general trends for distractor tones at more frequencies (total 8) relative to the A and B tones (Figure 4F). One experiment tested distractors aligned with the A (disA), the B (disB) or directly between (disAB). These conditions showed a significant effect on proportion segregated [ $F_{(3,21)} = 5.00, P < 0.05$ ]. Another experiment tested distractors above B (disB+4, disB+8) and below A (disA-2). These conditions did not show a significant effect [ $F_{(3,21)} = 2.145, P = 0.125$ ],



which is indicative of the diminishing effect of the distractors away from the A and B tones.

### 3. DISCUSSION

We report new insights on the dynamics of build-up in perceptual segregation, including the initial bias toward integration, and the effects of pauses, distractor tones and deviant tones. In audition the initial percept is typically integration with segregation developing over seconds (Anstis and Saida, 1985; Bregman, 1994). But such biasing toward integration has eluded neuronally-based explanation. We suggest that the initial bias is determined by broad onset activation in neurons selective to low-level features (e.g., tone frequency Micheyl et al., 2005; Pressnitzer et al., 2008) in or even below primary sensory cortex, prior to early adaptation and emergence of strong feature dependence. This property at onset biases the initial conditions that are propagated downstream of A1 to areas involved in identification of perceptual patterns and competition between them where build-up develops more slowly. Our study focused on auditory streaming, but the principle could generalize to motion plaid displays consisting of two gratings moving in different directions, also showing an initial bias toward integrated pattern motion (Hupé and Pressnitzer, 2012). Neural responses in visual areas representing the relevant low-level feature (motion direction) show a broader initial activation and a bias toward the vector average, i.e., integrated, direction (Recanzone and Wurtz, 1999). Our experiments and modeling demonstrate that the bias in the auditory case is partially restored during pauses that allow some recovery from early and fast adaptation (as brief as sub-sec), thereby allowing a refresh of the biased initial conditions. While various changes in stimuli can also interrupt build-up and reset the percept toward integration we have discovered a class of perturbations that promote segregation rather than integration. In auditory streaming a transient perturbation that disrupts the triplet pattern (e.g., replacing B by a deviant D or adding a distractor tone d between triplets) promotes segregation. Intuitively, these events could briefly make one of the streams more salient and cause a switch. Our model provided an opportunity to seek a more mechanistic explanation. Based on our experiments and modeling we propose that non-triplet tones are processed differently downstream from primary auditory cortex. Furthermore, our results support the notion of auditory streaming being bistable between perceptual states, where a pause or aberrant tone can flip the percept in a specific direction and the perturbation's effect is still evident several seconds later.

#### 3.1. Timescale of the Reset to Integrated

Using short stimulus presentations, we confirmed a partial reset to integrated for pauses of 300 or 600 ms, but did not find an increasing trend between the two conditions. A reset to integrated has been shown with pauses longer than 1 s using short stimulus presentations (Cusack et al., 2004; Snyder et al., 2008, 2009) and with briefer pauses (<1 s) using long stimulus presentations (during bistable alternations) (Denham

et al., 2010). Sussman et al. (2007) showed a reset for multi-second pauses, using EEG recordings and a mismatch negativity paradigm.

In our model, initial A1 responses had a large amplitude and broad tonotopic tuning; fast adaptation on a common timescale led to static responses with lower amplitude and tightened tuning. The tonotopic component is key for the initial bias toward integration. A rapid recovery of the fast adaptation following a stimulus pause led to a partial reset to integrated, consistent with our data. For a long enough pause there must be a full reset to integrated, as if hearing the stimulus for the first time. Bregman (1978) suggested biasing from previous stimuli would have recovered within 6 s and this was confirmed by later studies (Beauvois and Meddis, 1997; Snyder et al., 2008). Our results suggest that although rapid recovery of adaptation in A1 may explain the partial reset to integrated (even for very brief stimulus pauses), the multi-second timescale of longer term recovery may also be related to processes downstream of A1.

#### 3.2. Link between Context and Perturbations

A sudden change after a sequence of context triplets causes at least partial resetting of build-up back toward integration, as shown for a change in ear of presentation (Anstis and Saida, 1985), a shift in perceived loudness and/or location (Rogers and Bregman, 1993, 1998), a switch in attention (Carlyon et al., 2001; Thompson et al., 2011) and a pause in presentation (as described above); see review Moore and Gockel (2012). Like a pause, a switch in attention could allow recovery from adaptation. Otherwise a one-time shift of the entire stimulus in location or intensity (an increase, but not decrease Rogers and Bregman, 1998) could recruit previously unstimulated and, therefore, unadapted neurons. We may view the triplets preceding the deviant and distractor tones during build-up as context. Different types of context can bias perception toward (i.e., prime for) integration or segregation (Sussman and Steinschneider, 2006; Snyder et al., 2008, 2009; Rahne and Sussman, 2009). Even for a context of a single stream of tones, say A\_A\_A\_A\_, that would alone promote segregation for subsequent test triplets ABA\_ ABA\_..., similar disruptions as above at the end of the context sequence lead to integration, as if the effect of the context was undone (Rogers and Bregman, 1993, 1998; Beauvois and Meddis, 1997). Also, a single deviant A' at the end of an A\_A\_A\_...context can reduce or eliminate the expected bias toward segregation (Roberts et al., 2008; Haywood and Roberts, 2010, 2013). So while these various disruptions favor integration and may a priori lead one to a generalized expectation that a single transient distractor tone (between triplets) or a single deviant tone (within a triplet) should also promote integration, we found the opposite — promotion of segregation in the subsequent triplets. Nevertheless our results do not contradict these previous studies. Studies looking at the effects of deviant tones did so by placing these at the end of a single stream context (Roberts et al., 2008; Haywood and Roberts, 2010, 2013); in our study we placed the deviant or distractor at the end of context triplets. Thompson et al. (2011) included an experiment

with a single delayed-onset deviant B tone, but did not report promotion of segregation or resetting. Still, other stimulus perturbations may promote segregation. Further experiments should consider whether single-tone deviants in features other than frequency (e.g., lateralization or loudness) can promote segregation.

### 3.3. Promotion of Segregation and Differential Processing of Non-triplet Events

Our model accounts for the observed segregation-promoting effects by assuming that inputs propagate to the competition stage in a differential manner: A1 responses to a deviant or a distractor tone do not reach the “integration (AB) unit.” It encodes a non-trivial rhythm and can be viewed as more sensitive to, effectively selective against, sounds that break the triplet pattern. Our implicit assumption is that the aberrant tone is identified as a mismatch and is deflected from reaching AB. Viewed differently, an incoming sound inconsistent with the integrated percept might result in the integration unit being briefly suppressed, allowing the peripheral units to take over. The crucial aspect is that the incoming tones have a differential effect on the integrated and segregated units. The effects of distractor tones also show a dependence on tonotopy, which led us to favor an input-based explanation.

Our results allow us to rule out some other potential explanations for the effects of non-triplet perturbations. Suppose that such perturbations indiscriminately cause a switch in perception away from the current percept. One might argue that we saw switches only from integrated to segregated since we considered perturbations only during build-up, when integration is thought to be dominant. However, our data do not support the idea of switches from segregated to integrated. At DF=10, where ~50% of trials are already segregated after 3 triplets, we saw no evidence of a reset or switch back toward integration (**Figure 2E**), either in the group data, or in individual subjects (not shown). This facet of the data is consistent with the proposed notion that input propagates from A1 to the segregated units, but not the integrated unit. Another hypothesis could be that any transient, salient perturbation distinct from standard triplets promotes segregation. However, our data showed no effect for distractor tones sufficiently far in frequency from the A or B tones. Our modeling work shows that this interaction could be through input from the distractor tones still propagating to segregated units with tonotopic dependence.

Haywood and Roberts (2013) showed that hearing a single A tone before the triplets could make that stream more salient. Could there be a similar effect in our data, where the distractor tone makes one of the streams more salient, or briefly directs attention toward it? The range of conditions for which we found promotion of segregation includes several cases where the perturbation is not an A or a B tone. The distractor tone *d* appears in a sequence ...ABA\_ABAdABA.... It could be that the *d* is being grouped into a new triplet (AdA or B-d-B), thus making the A or B stream more salient (or highlighting their separation) ahead of the upcoming test triplets. For a distractor

or deviant tone, the proposed mechanism in our model boosts inputs to the competition stage for the segregated units whilst gating out input to the integrated unit. This selective transient modulation of input gains could be viewed as a brief top-down attentional effect. However, for an attention mechanism, the selective gain would likely act in response to the perturbation mismatch with some delay. In our current model we have idealized the transmission of input from A1 to competition stage without a delay.

### 3.4. Build-Up and Bistability in Models

Most existing computational models of auditory streaming have focused on reproducing the dependence (van Noorden, 1975) of perceptual bias on DF and presentation rate (Almonte et al., 2005; Elhilali and Shamma, 2008; Wang and Chang, 2008), the dynamics of build-up (Beauvois and Meddis, 1991; McCabe and Denham, 1996) or both (Beauvois and Meddis, 1996). A complete theoretical framework for streaming should account for build-up as well as the later alternations, given that the probability of perceiving segregation converges to the long-term probability of bistable alternations. Some recent models focused on post-build-up alternations (auditory bistability) (Mill et al., 2013; Barniv and Nelken, 2015; Rankin et al., 2015; Steele et al., 2015). The initial bias to integration is set by specifying a priori initial conditions (Barniv and Nelken, 2015; Steele et al., 2015). In Mill et al. (2013) the bias is emergent through an early stage of an algorithmic pattern discovery. Our model that accounts for alternations, and was further developed here to describe build-up, is the first treatment to explain the initial bias for integration through a direct link to observed neurophysiological responses (Fishman et al., 2001; Micheyl et al., 2005). To the best of our knowledge, no other model has been used to investigate resetting effects, or the effects of perturbations in general.

### 3.5. Future Directions for Our Model

Our current neuromechanistic model relies on a lumped version of a distributed network, where a few discrete units pool inputs from different tonotopic locations in A1. Although this view allows the model to account for many phenomena (stimulus parameter dependence, build-up, alternations, resetting for pauses), the notion of differential processing introduced to account for promotion of segregation approaches the limit of our modeling framework, and suggests the need for a richer description. One avenue for extension would be to consider a continuous feature space in DF, as proposed in Almonte et al. (2005), at least at the A1 stage of the model. Although the rules for the tonotopic spread of A1 responses allowed us to consider, for example, distractor tones away from the three locations A, B and  $(A+B)/2$ , a more refined description would define how A1 responds in time to any combination of tones across DF (and consider other paradigms, for example, involving frequency-banded maskers Elhilali et al., 2009b). As a further extension we could introduce an additional dimension to the feature space, e.g., selectivity to different repetition rates of the streams and the relative timing (phase) of the inputs. A first step in this direction was made in Beauvois and Meddis (1996),



with the use of delay lines to introduce a temporal feature space. Beyond this, a suitable theoretical framework to study might be a coupled oscillator network sensitive to frequencies in the range of the repetition rates of the tone sequences, not just the frequency of the tones (like tonotopy in A1). fMRI studies have implicated a broader network involved in streaming, including areas associated with rhythm and timing (Kashino and Kondo, 2012). The design of such a network and the necessary mechanisms for competition could build directly on our present model. Such networks have been used in studies of rhythm perception (Large et al., 2015) and in phenomenological studies of perceptual bistability (Borisjuk et al., 2009). Such a richer description would allow one to pursue the origin of the differential processing we propose here and to investigate the effects of temporal coherence, a strong cue in auditory stream segregation (Elhilali et al., 2009a).

## 4. CONCLUSION

Our model with the developments presented here is the first grounded in neurophysiological detail to account for build-up and subsequent bistable alternations. We propose that the initial bias to integrated arises naturally from the rapid but delayed emergence of low-level feature dependence and that the more gradual timescale of build-up comes from competition mechanisms downstream of A1. This is the first explanation of integration bias and build-up motivated directly from neurophysiological data (responses to triplet sequences in A1 Micheyl et al., 2005).

New findings presented here challenge the current understanding of how the segregation of auditory objects is affected by interruptions and perturbations. A reset of the build-up process results from an established class of perturbations that shift the entire triplet stimulus in location, loudness or timing. We illustrate that the rapid recovery of responses in A1 can explain resetting for stimulus pauses. We demonstrated a new and opposite effect, promotion of segregation, by a complementary class of perturbations that transiently alter a single triplet or introduce a new non-triplet element. Our modeling in conjunction with confirmed experimental predictions led to a new hypothesis: that new non-triplet events (deviant or distractor tones) are gated out from the neural population encoding the complex integrated rhythm.

## 5. MATERIALS AND METHODS

### 5.1. Neuromechanistic Model

The neuronal circuits for competition and perceptual encoding are assumed to be downstream and receiving inputs from A1. The periodic inputs mimic the A1-responses to ABA\_ sequences reported in Micheyl et al. (2005). Neuronal activity is described by mean firing rates and competitive interactions emerge through a combination of excitatory and inhibitory connections, slow adaptation and intrinsic noise. We provide a brief outline of the model architecture, mechanisms and inputs here; the full model

equations and further details can be found in Appendix B in Supplementary Material.

The schematic in **Figure 1A** shows downstream units *A*, *B*, and *AB* that respectively pool inputs from regions of A1 centered at locations with best frequencies *A*, *B* and the midpoint between  $(A + B)/2$ . We associate a variable  $r_k$  ( $k = \{A, AB, B\}$ ) with each unit representing the mean firing rate of a population of neurons in the competition network. For each unit  $r_k$  the intrinsic dynamics are illustrated in **Figure 1A** and described by a differential equation like the following,

$$\tau_r \dot{r}_{AB} = -r_{AB} + F\left(\beta_e e_{AB} - \beta_i (r_{AB} + r_A + r_B) - g a_{AB} + I_{AB} + \chi_{AB}\right). \quad (1)$$

By way of an example, we describe this equation for  $r_{AB}$  in detail; the equations for  $r_A$  and  $r_B$  take the same general form. The cortical timescale is  $\tau_r = 10$  ms. A sigmoidal firing rate (smooth threshold) function  $F$  (see Appendix B in Supplementary Material) process all inputs to the unit. Local excitation  $e_{AB}$  has strength  $\beta_e = 0.65$  and evolves on an NMDA-like timescale  $\tau_e = 70$  ms. Global inhibition (assumed instantaneous and so proportional to the cortical variables  $r_k$ ) has fixed strength  $\beta_i = 0.3$  independent of DF. Note  $\beta_e > \beta_i$  so there is net local excitation. Linear spike frequency adaptation (slow negative feedback)  $a_{AB}$  has strength  $g = 0.045$  and a timescale of 1.4 s.

The input  $I_{AB}$  mimics A1 cortical responses to triplet tone sequences; full details are given in Appendix B in Supplementary Material. There are two components to the early adaptation of these responses, both consistent with observations from Micheyl et al. (2005) and sharing a common timescale  $\tau_{A1} = 500$  ms (**Figure 1B**). Firstly, the overall amplitude of responses decays. Secondly, the effective DF is initially small i.e. the DF dependence of the responses takes time to emerge. After a stimulus pause, the A1 adaptation is assumed to rapidly recover ( $\tau_{rec} = 100$  ms), such that when the stimulus resumes after an adequate pause (say  $2 \times \tau_{rec}$ ) the model inputs resemble those after initial stimulus onset (**Figure 3B**). For a distractor tone (or a deviant tone) input amplitude and tonotopic spread are consistent with a partially recovered response to the tone. At the tonotopic location *A*, responses to a distractor are reduced, because the distractor immediately follows the offset of a normal *A* tone (referred to as temporal forward masking in Fishman et al., 2001). Intrinsic additive noise  $\chi_{AB}$  is an independent Ornstein-Uhlenbeck process for each  $r_k$ .

Numerical simulations were implemented with an Euler-Murayama scheme with a stepsize of  $0.5\tau_r$ . Build-up functions were computed as time-binned averages across 500 simulations. For each time bin the fraction of trials with more activity in the *AB* unit than the summed activities of the *A* and *B* units was taken as the measure of proportion integrated. Computations were implemented in `Matlab` and batch processed using the function `parfor`; no special computing hardware was required. In all computations, the same set of 500 randomized initial conditions and the same 500 instantiations of the noise process (i.e., frozen noise) was used for each  $r_k$ . This ensures that any

differences between conditions is entirely due to changes to model parameters (e.g., reflecting different stimulus properties). For example, in **Figures 3B,E**, the control (No pause, No dis) curves only deviate from the test simulations (Pause 600 ms, disB+2) from the time point where the perturbation is introduced.

## 5.2. Psychoacoustic Experiments

Our experimental paradigm is well suited for studying the effects of perturbations on how the subsequent triplets are perceived. In all experiments presented here (with pauses, distractors or deviants) the perturbation was followed by three normal triplets and subjects reported their perception of the final triplet, roughly 1 s after the perturbation. Three triplets provides enough stimulus duration to make a reliable perceptual judgment (Carlyon et al., 2003; Haywood and Roberts, 2010). This design precludes the possibility of subjects reporting, say, a distractor tone as being its own segregated stream, as the distractor occurs well before the final triplet. If continuous perceptual reports were used, confusion might arise about classifying an unexpected tone into its own stream at the moment the distractor is detected. A final possibility would be to use an objective measure of streaming (Roberts et al., 2008; Thompson et al., 2011). An appropriate paradigm could be the one used in Thompson et al. (2011), where performance in a deviant detection task functioned as an objective measure for streaming and showed qualitative agreement with subjective perceptual reports. In the objective task, subjects had to detect a single delayed-onset B tone and performance was best during integration. Given that the objective task relies on the detection of a delayed-onset deviant and that some trials would need to involve another deviant tone (the perturbations studied here), it could become rather confusing for a subject. It would be challenging for a subject to distinguish between multiple types of aberrant tone, ignoring some and reporting the presence of others.

### 5.2.1. Procedure

Subjects sat in an acoustically shielded chamber and pressed keys on a keyboard to indicate their perceptual response. In each task, a short ABA<sub>1</sub> sequence ranging between three and 10 triplets was played, and the subjects reported with button presses whether the last triplet of the sequence sounded most like the integrated percept or the segregated percept and guessed if unsure. The integrated percept was defined as hearing the A and the B tones together in a galloping rhythm, and the segregated percept was defined as hearing the A tones and B tones separately in two distinct streams. Subjects were instructed to respond as quickly as possible and had up to 5 s—the length of the inter-stimulus interval (ISI)—to respond.

### 5.2.2. Stimuli

The repeating ABA<sub>1</sub> triplet consist of 100 ms pure tones with 10 ms linear ramps, where the “\_” indicates a silence also lasting 100 ms; in total, the duration of each ABA<sub>1</sub> triplet is 400 ms. An inter-stimulus interval of 5 s was included between all trials. The higher frequency B tones are a variable DF semitones (st) above the lower frequency A tones. Each tone sequence was played

binaurally through Etymotic headphones at 65 dB SPL. Three DF conditions were used for all experiments:  $DF \in \{4, 7, 10\}$  st. From trial to trial the A-tone base-frequencies were roved between 420 and 1060 Hz, separated by intervals of 4 st; correspondingly, the B tone frequencies ranged between 530 and 1888 Hz. The roving of base frequencies and ISI of 5 s were chosen to avoid any latent adaptation from one trial to the next (Beauvois and Meddis, 1997; Snyder et al., 2009; Sussman-Fort and Sussman, 2014).

### 5.2.3. Subjects

Seventeen subjects in total, including one of the authors, took part in the experiments (10 female, 7 male), aged 20–51, mean age 27.9. Subjects were reimbursed for their participation and all experimental procedures complied with human subject research guidelines as approved by the University Committee on Activities Involving Human Subjects at New York University (IRB-FY2016-310). All subjects provided written informed consent and were required to pass a hearing screening.

### 5.2.4. Conditions

The stimulus paradigm for the pause experiment is shown in **Figure 2A**. A total of 15 conditions (3 DF conditions crossed with 5 stimulus length/pause combinations) were tested with 20 repetitions of each condition (total of 300 trials for each of 8 subjects). Two test conditions consisted of 7 context triplets followed by a pause of 300 or 600 ms, which was then followed by 3 test triplets. Three control conditions of 3 (test only), 7 (context only), and 10 (no pause control) triplets were tested in 9 blocks of 20 trials and the test conditions in 8 blocks of 15 trials. Control and test conditions were run in separate block sections to avoid confusion about timing of perceptual reports.

Schematics of the stimulus paradigm for the distractor and deviant experiments are shown in **Figures 2B,C**. Three different experimental sessions, with eight subjects each, were conducted for different experimental conditions. Subjects performed 20 blocks of 15 trials each, where the length of each trial ranged from 1.2 s to 2.4 s in length. In each experiment, two control conditions included a 3 triplet and a 6 triplet condition with no deviants or distractors. Along with the two control conditions, each experiment included three distractor or deviant conditions, 6 triplets in length. Distractor tones were 50 ms in length and were inserted symmetrically in the 100ms inter-triplet gap between the third and fourth triplets of the sequence, so that there was 25ms of silence on either side of the distractor. Across the three experimental sessions, the following frequencies (in st, relative to the A and B tones of the triplets) of distractor tones were tested: A–2, A, (A+B)/2, B, B+2, B+4, B+8, B+15. Deviant tones involved a change in frequency to the B tone of the third triplet. In the one deviant tone condition tested, the B-tone was increased in frequency by 2 st.

## AUTHOR CONTRIBUTIONS

JRa and PO designed psychoacoustic experiments. PO carried out psychoacoustic experiments. JRa and PO analyzed

psychoacoustic data. JRa and JRi conceived the computational model. JRa did modeling simulations. JRa and JRi wrote the manuscript.

## ACKNOWLEDGMENTS

The authors express their gratitude to Rodica Curtu, Yonatan Fishman, David Heeger, Antje Ihlefeld, Andrew Oxenham and Elyse Sussman for helpful feedback on earlier versions of this

manuscript. JRa was supported by a Swartz Foundation postdoc grant. PO received support from the NYU Dean's Undergraduate Research Fund. Partial support for JRi comes from an NIH K-18 award, K18 DC011602.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fnins.2017.00198/full#supplementary-material>

## REFERENCES

- Almonte, F., Jirsa, V., Large, E., and Tuller, B. (2005). Integration and segregation in auditory streaming. *Phys. D Nonlinear Phenomena* 212, 137–159. doi: 10.1016/j.physd.2005.09.014
- Anstis, S., and Saida, S. (1985). Adaptation to auditory streaming of frequency-modulated tones. *J. Exp. Psychol. Hum. Percept. Perform.* 11, 257–271. doi: 10.1037/0096-1523.11.3.257
- Barniv, D., and Nelken, I. (2015). Auditory streaming as an online classification process with evidence accumulation. *PLoS ONE* 10:e0144788. doi: 10.1371/journal.pone.0144788
- Beauvois, M., and Meddis, R. (1991). A computer model of auditory stream segregation. *Q. J. Exp. Psychol.* 43, 517–541. doi: 10.1080/14640749108400985
- Beauvois, M., and Meddis, R. (1996). Computer simulation of auditory stream segregation in alternating-tone sequences. *J. Acoust. Soc. Am.* 99, 2270–2280.
- Beauvois, M., and Meddis, R. (1997). Time decay of auditory stream biasing. *Percept. Psychophys.* 59, 81–86.
- Bee, M., and Klump, G. (2004). Primitive auditory stream segregation: a neurophysiological study in the songbird forebrain. *J. Neurophysiol.* 92, 1088–1104. doi: 10.1152/jn.00884.2003
- Bee, M., Micheyl, C., Oxenham, A., and Klump, G. (2010). Neural adaptation to tone sequences in the songbird forebrain: patterns, determinants, and relation to the build-up of auditory streaming. *J. Comp. Physiol. A* 196, 543–557. doi: 10.1007/s00359-010-0542-4
- Borisjuk, R., Chik, D., and Kazanovich, Y. (2009). Visual perception of ambiguous figures: synchronization based neural models. *Biol. Cyber.* 100, 491–504. doi: 10.1007/s00422-009-0301-1
- Bregman, A. (1978). Auditory streaming is cumulative. *J. Exp. Psychol. Hum. Percept. Perform.* 4:380.
- Bregman, A. (1994). *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: MIT Press.
- Carlyon, R., Cusack, R., Foxton, J., and Robertson, I. (2001). Effects of attention and unilateral neglect on auditory stream segregation. *J. Exp. Psychol. Hum. Percept. Perform.* 27, 115–127. doi: 10.1037/0096-1523.27.1.115
- Carlyon, R., Plack, C., Fantini, D., and Cusack, R. (2003). Cross-modal and non-sensory influences on auditory streaming. *Perception* 32, 1393–1402. doi: 10.1068/p5035
- Cusack, R., Decks, J., Aikman, G., and Carlyon, R. (2004). Effects of location, frequency region, and time course of selective attention on auditory scene analysis. *J. Exp. Psychol. Hum. Percept. Perform.* 30:643. doi: 10.1037/0096-1523.30.4.643
- Denham, S., Gyimesi, K., Stefanics, G., and Winkler, I. (2010). “Stability of perceptual organization in auditory streaming,” in *The Neurophysiological Bases of Auditory Perception*, eds E. Lopez-Poveda, A. R. Palmer, and R. Meddis (New York, NY: Springer), 477–487.
- Denham, S., Gyimesi, K., Stefanics, G., and Winkler, I. (2013). Perceptual bistability in auditory streaming: how much do stimulus features matter? *Learn. Percept.* 5, 73–100. doi: 10.1556/LP.5.2013.Supp1.6
- Elhilali, M., Ma, L., Micheyl, C., Oxenham, A., and Shamma, S. (2009a). Temporal coherence in the perceptual organization and cortical representation of auditory scenes. *Neuron* 61, 317–329. doi: 10.1016/j.neuron.2008.12.005
- Elhilali, M., and Shamma, S. (2008). A cocktail party with a cortical twist: how cortical mechanisms contribute to sound segregation. *J. Acoust. Soc. Am.* 124, 3751–3771. doi: 10.1121/1.3001672
- Elhilali, M., Xiang, J., Shamma, S., and Simon, J. (2009b). Interaction between attention and bottom-up saliency mediates the representation of foreground and background in an auditory scene. *PLoS Biol.* 7:e1000129. doi: 10.1371/journal.pbio.1000129
- Fishman, Y., Arezzo, J., and Steinschneider, M. (2004). Auditory stream segregation in monkey auditory cortex: effects of frequency separation, presentation rate, and tone duration. *J. Acoust. Soc. Am.* 116, 1656–1670. doi: 10.1121/1.1778903
- Fishman, Y., Reser, D., Arezzo, J., and Steinschneider, M. (2001). Neural correlates of auditory stream segregation in primary auditory cortex of the awake monkey. *Hear. Res.* 151, 167–187. doi: 10.1016/S0378-5955(00)00224-0
- Haywood, N., and Roberts, B. (2010). Build-up of the tendency to segregate auditory streams: resetting effects evoked by a single deviant tone. *J. Acoust. Soc. Am.* 128, 3019–3031. doi: 10.1121/1.3488675
- Haywood, N., and Roberts, B. (2013). Build-up of auditory stream segregation induced by tone sequences of constant or alternating frequency and the resetting effects of single deviants. *J. Exp. Psychol. Hum. Percept. Perform.* 39, 1652–1666. doi: 10.1037/a0032562
- Hupé, J.-M., and Pressnitzer, D. (2012). The initial phase of auditory and visual scene analysis. *Philos. Trans. R. Soc. B Biol. Sci.* 367, 942–953. doi: 10.1098/rstb.2011.0368
- Itatani, N., and Klump, G. (2014). Neural correlates of auditory streaming in an objective behavioral task. *Proc. Natl. Acad. Sci. U.S.A.* 111, 10738–10743. doi: 10.1073/pnas.1321487111
- Kashino, M., and Kondo, H. (2012). Functional brain networks underlying perceptual switching: auditory streaming and verbal transformations. *Philos. Trans. R. Soc. B Biol. Sci.* 367, 977–987. doi: 10.1098/rstb.2011.0370
- Large, E., Herrera, J., and Velasco, M. (2015). Neural networks for beat perception in musical rhythm. *Front. Sys. Neurosci.* 9:159. doi: 10.3389/fnsys.2015.00159
- McCabe, S., and Denham, M. (1996). A model of auditory streaming. *Adv. Neural Inform. Process. Syst.* 10, 1611–1621.
- Micheyl, C., Tian, B., Carlyon, R., and Rauschecker, J. (2005). Perceptual organization of tone sequences in the auditory cortex of awake macaques. *Neuron* 48, 139–148. doi: 10.1016/j.neuron.2005.08.039
- Mill, R., Böhm, T., Bendixen, A., Winkler, I., and Denham, S. (2013). Modelling the emergence and dynamics of perceptual organisation in auditory streaming. *PLoS Comput. Biol.* 9:e1002925. doi: 10.1371/journal.pcbi.1002925
- Moore, B., and Gockel, H. (2012). Properties of auditory stream formation. *Philos. Trans. R. Soc. B Biol. Sci.* 367, 919–931. doi: 10.1098/rstb.2011.0355
- Moreno-Bote, R., Rinzal, J., and Rubin, N. (2007). Noise-induced alternations in an attractor network model of perceptual bistability. *J. Neurophysiol.* 98, 1125–1139. doi: 10.1152/jn.00116.2007
- Pressnitzer, D., and Hupé, J. (2006). Temporal dynamics of auditory and visual bistability reveal common principles of perceptual organization. *Curr. Biol.* 16, 1351–1357. doi: 10.1016/j.cub.2006.05.054
- Pressnitzer, D., Sayles, M., Micheyl, C., and Winter, I. (2008). Perceptual organization of sound begins in the auditory periphery. *Curr. Biol.* 18, 1124–1128. doi: 10.1016/j.cub.2008.06.053

- Rahne, T., and Sussman, E. (2009). Neural representations of auditory input accommodate to the context in a dynamically changing acoustic environment. *Eur. J. Neurosci.* 29, 205–211. doi: 10.1111/j.1460-9568.2008.06561.x
- Rankin, J., Sussman, E., and Rinzel, J. (2015). Neuromechanistic model of auditory bistability. *PLoS Comput. Biol.* 11:e1004555. doi: 10.1371/journal.pcbi.1004555
- Recanzone, G., and Wurtz, R. (1999). Shift in smooth pursuit initiation and MT and MST neuronal activity under different stimulus conditions. *J. Neurophysiol.* 82, 1710–1727.
- Roberts, B., Glasberg, B., and Moore, B. (2008). Effects of the build-up and resetting of auditory stream segregation on temporal discrimination. *J. Exp. Psychol. Hum. Percept. Perform.* 34:992. doi: 10.1037/0096-1523.34.4.992
- Rogers, W., and Bregman, A. (1993). An experimental evaluation of three theories of auditory stream segregation. *Percept. Psychophys.* 53, 179–189.
- Rogers, W., and Bregman, A. (1998). Cumulation of the tendency to segregate auditory streams: resetting by changes in location and loudness. *Percept. Psychophys.* 60, 1216–1227. doi: 10.3758/BF03206171
- Scholes, C., Palmer, A., and Sumner, C. (2015). Stream segregation in the anesthetized auditory cortex. *Hear. Res.* 328, 48–58. doi: 10.1016/j.heares.2015.07.004
- Shapiro, A., Moreno-Bote, R., Rubin, N., and Rinzel, J. (2009). Balance between noise and adaptation in competition models of perceptual bistability. *J. Comput. Neurosci.* 27, 37–54. doi: 10.1007/s10827-008-0125-3
- Snyder, J., Carter, O., Lee, S.-K., Hannon, E., and Alain, C. (2008). Effects of context on auditory stream segregation. *J. Exp. Psychol. Hum. Percept. Perform.* 34:1007. doi: 10.1037/0096-1523.34.4.1007
- Snyder, J. S., Carter, O. L., Hannon, E. E., and Alain, C. (2009). Adaptation reveals multiple levels of representation in auditory stream segregation. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 1232–1244. doi: 10.1037/a0012741
- Steele, S., Tranchina, D., and Rinzel, J. (2015). An alternating renewal process describes the buildup of perceptual segregation. *Front. Comput. Neurosci.* 8:166. doi: 10.3389/fncom.2014.00166
- Sussman, E., Horváth, J., Winkler, I., and Orr, M. (2007). The role of attention in the formation of auditory streams. *Percept. Psychophys.* 69, 136–152. doi: 10.3758/BF03194460
- Sussman, E., and Steinschneider, M. (2006). Neurophysiological evidence for context-dependent encoding of sensory input in human auditory cortex. *Brain Res.* 1075, 165–174. doi: 10.1016/j.brainres.2005.12.074
- Sussman-Fort, J., and Sussman, E. (2014). The effect of stimulus context on the buildup to stream segregation. *Front. Neurosci.* 8:93. doi: 10.3389/fnins.2014.00093
- Taaseh, N., Yaron, A., and Nelken, I. (2011). Stimulus-specific adaptation and deviance detection in the rat auditory cortex. *PLoS ONE* 6:e23369. doi: 10.1371/journal.pone.0023369
- Thompson, S., Carlyon, R., and Cusack, R. (2011). An objective measurement of the build-up of auditory streaming and of its modulation by attention. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 1253–1262. doi: 10.1037/a0021925
- Ulanovsky, N., Las, L., Farkas, D., and Nelken, I. (2004). Multiple time scales of adaptation in auditory cortex neurons. *J. Neurosci.* 24, 10440–10453. doi: 10.1523/JNEUROSCI.1905-04.2004
- van Noorden, L. (1975). *Temporal Coherence in the Perception of Tone Sequences*. Ph.D. thesis, Eindhoven University of Technology.
- Wang, D., and Chang, P. (2008). An oscillatory correlation model of auditory streaming. *Cogn. Neurodyn.* 2, 7–19. doi: 10.1007/s11571-007-9035-8

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Rankin, Osborn Popp and Rinzel. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.