Brainstem mechanisms contributing to accurate source localisation are suited to the reverberant structure of natural acoustic scenes

Andrew Brughera[1,2,4], Jason Mikiel-Hunter[1,4], Mathias Dietz[3], & David McAlpine[1]

[1] Department of Linguistics, and the Australian Hearing Hub, Macquarie University, 16 University Avenue, Macquarie University, NSW, 2109, Australia

[2] Department of Biomedical Engineering, Boston University, 44 Cummington Mall, Boston, MA 02215, U.S.A.

[3] Medizinische Physik and Cluster of Excellence "Hearing4all" Universität Oldenburg, 26111 Oldenburg, Germany

[4] A.B. and J.M.H. contributed equally to this work.

ABSTRACT

Listeners perceive sound energy as originating from the direction of its source, even when followed only milliseconds later by reflected energy from multiple different directions. Here, modelling responses of brainstem neurons responsible for encoding auditory spatial cues, we demonstrate that accurate localisation in reverberant environments relies on pre- and post-synaptic specializations that emphasise spatial information in early-arriving sound energy. Synaptic depression in spherical bushy cells (SBCs) of the cochlear nucleus, and outwardly-rectifying membrane currents in SBCs and neurons of the medial superior olive (MSO) to which SBCs project bilaterally, are necessary to account for human listening performance. These biophysical properties appear suited to efficient coding of spatial information, emphasising early-arriving spatial information, particularly at sound frequencies where reverberant energy is relatively intense. Applied to lateralisation of human speech in a virtual reverberant room, we show that frequency-dependent membrane properties enhance correct, over spurious, localisation cues at the earliest stages of spatial processing.

INTRODUCTION

In many listening environments, sound propagating directly from its source to a listener's ears arrives first and more intensely at the near ear, microseconds later and less intensely at the far ear, and is then followed milliseconds later by multiple reverberant copies arriving from different directions. Despite this mixture of direct and reflected sound pressure that generates non-stationary and rapidly fluctuating spatial cues in the interaural time difference (ITD) and interaural intensity difference (IID), listeners typically perceive a sound as punctate, and originating from the direction of its source. This ability to determine the true location of the source persists even when the intensity of reflected sound matches, or even exceeds, that of the early-arriving, direct sound (Brown et al., 2015; Dietz et al., 2013; Haas, 1951; Hartung & Trahiotis, 2001; Hu et al., 2017; Litovsky et al., 1999; Wallach et al., 1949), contributing to phenomena such

as ‘cocktail party listening’ – attending to a single talker against a background of many competing voices (Cherry, 1953).

Neural correlates of ‘echo-suppression’ have been reported along the ascending auditory pathway, from brainstem (Dietz et al., 2014), to midbrain (Fitzpatrick et al., 1995; Litovsky & Delgutte, 2002; Litovsky & Yin, 1998a,b; Tollin et al., 2004; Yin, 1994), and primary auditory cortex (Liebenthal & Pratt, 2002). In particular, brainstem neurons of the medial superior olive (MSO)—which encode spatial information through their sensitivity to ITDs conveyed in the temporal fine structure (TFS) of low-frequency (< 1500 Hz) sounds (Mathews et al., 2010; Scott et al., 2010; Yin & Chan, 1990)—respond strongly to rising sound-energy that would emphasise early-arriving sound direct from the source (Dietz et al., 2014). Monaural projection pathways to binaural MSO neurons appear critical to this emphasis of direct sound: modelling of physiological and perceptual data suggests that neural adaptation arising before binaural interaction is necessary to emphasise TFS ITDs conveyed in rising sound-energy (Dietz et al., 2013, 2014).

In many outdoor environments, including urban streets and rural forests and fields, reverberation-time (‘T60’—the time for reverberant energy to fall by 60 decibels) systematically decreases as sound frequency decreases below 1500 Hz (Traer & McDermott, 2016), leading to sounds reaching the ears with a generally higher direct-to-reverberant ratio (DRR). Brain mechanisms that efficiently transduce auditory spatial cues (Harper & McAlpine, 2004; Harper et al., 2014) might be expected to reflect the decreasing influence of reverberation at relatively lower sound-frequencies, which convey information critical to communication, such as the fundamental frequency of speech and other vocalisations, the harmonics of voice pitches, and the first formants of vowels. Consistent with this hypothesis, the behavioural emphasis of early-arriving ITD information in the TFS of low-frequency sounds is frequency dependent (Hu et al., 2017); listeners are more sensitive to ITDs conveyed in the rising-energy portion of amplitude-modulated 600-Hz sounds than they are to ITDs at the energy peak. This is not the case at 200 Hz, where listeners are equally sensitive to ITDs conveyed in rising and peak energy, and where, in natural settings, reverberant energy can be >20 dB less intense at 200 than at 600 Hz (Traer & McDermott, 2016). For both frequencies, listeners are least sensitive to ITD during falling sound-energy, when direct sound is most likely to be corrupted by reverberation. These data suggest that brain mechanisms that transmit reliable, and suppress unreliable, auditory-spatial information, do so accounting for natural frequency-profiles in reverberant energy across the range of sound frequencies at which TFS ITDs are biophysically encoded to determine source location.

Here, employing a standard model of the auditory periphery (Zilany et al., 2014) to drive a nonlinear brainstem model, we demonstrate two suppressive mechanisms to account for the frequency-dependent emphasis of rising and peak sound-energy in ITD encoding: (1) synaptic depression in spherical bushy cells (SBCs) of the ventral

cochlear nucleus (VCN) (Oleskevich et al., 2000; Wang & Manis, 2008; Yang & Xu-Friedman, 2009) that project excitatory inputs to binaural MSO neurons; and (2) outwardly-rectifying low-threshold potassium (K_{LT}, K_{v1}) channels, present in the MSO neurons themselves (Gai et al., 2009; Mathews et al., 2010; Svirskis et al., 2002, 2004), and in the SBCs (Manis & Marx, 1991; Rothman & Manis, 2003a,b,c). We show the importance of these features in accounting for sensitivity to ITDs conveyed in the rising energy of sounds at the level of single MSO neurons (Dietz et al., 2014). Then, with our nonlinear model, and a linear spiking model (Remme et al., 2014) extended to populations of MSO neurons in 'left' and 'right' brain hemispheres, we demonstrate brainstem properties consistent with the stronger weighting of ITD information conveyed in rising energy at 600 Hz—a frequency at which reflected sounds are likely to be more energetic—compared to 200 Hz, where reflected energy is relatively less intrusive.

Finally, employing our population model of both brain hemispheres, we relate the relative weighting of spatial information in rising energy to the correct lateralisation of natural speech in realistic acoustic environments ("Haas effect"; Haas, 1951), demonstrating that monaural synaptic depression prior to binaural interaction is essential to judging unambiguously the laterality of natural speech in a simulated reverberant space with direct sound and early reflections off virtual walls (Dietz et al., 2014). In particular, the dominance of ITD information conveyed in rising energy in the 600-Hz channel facilitates lateralisation of speech sounds by emphasising ITD cues of the direct source over the confounding ITD cues of its reflected copies.

Our physiologically realistic model of brainstem neurons can account for the strong suppression of reverberation for sound frequencies from 500 to 1200 Hz, a frequency range that despite being high in reverberation (Traer & McDermott, 2016) produces the lowest perceptual thresholds for ITD discrimination in human listeners (Brughera et al., 2013; Klumpp & Eady, 1956), and dominates the percept of auditory spatial cues (Ihlefeld and Shinn-Cunningham, 2011; Shinn-Cunningham et al., 1995; Xia et al., 2010).

RESULTS

Modelling frequency-dependent emphasis of auditory-spatial cues in rising and peak sound-energy

We first explored the extent to which synaptic depression at model SBCs in each cochlear nucleus accounts for the emphasis of ITDs in the rising energy of modulated sounds in a model MSO neuron (Fig. 2B). We did so, employing amplitude-modulated binaural beats [AMBBs; (Dietz et al., 2013, 2014)], in which binaurally presented tones of a slightly different frequency are modulated in amplitude at a rate equal to the difference in frequency across the two ears. This generates a stimulus in which each cycle of rising and falling sound energy contains a full cycle of interaural phase disparities. AMBBs (Fig. 2A) centred at 600 Hz, with modulation rates 4 to 64 Hz, were presented to a standard model of peripheral auditory processing (Zilany et al., 2014),

which drove a nonlinear model of brainstem neurons. The model of peripheral auditory processing includes 24 adapting auditory nerve (AN) fibres: twelve each in the left and right ears, with characteristic frequency (CF)[1] 600 Hz. Each of three distinct model AN fibres projects an ipsilateral depressing, excitatory synaptic input (Rudnicki & Hemmert, 2017; u = 0.5) to one of the eight (four left, and four right) model SBCs. Each SBC projects a non-depressing, excitatory synaptic input to the multi-compartment model MSO neuron. Both the model SBCs and MSO neuron are Hodgkin & Huxley (1952) models, with circuit elements that represent membrane ion-channels, including K_{LT} (Mathews et al., 2010; Rothman & Manis, 2003c; Scott et al., 2010).

Beginning without synaptic depression, a nonlinear model with very strong, non-depressing, supra-threshold synapses in each cochlear nucleus [developed to generate strong, sustained responses (Joris et al., 1994) and ITD sensitivity to unmodulated pure-tone stimuli (Yin & Chan, 1990)] produced only slight adaptation in the spike rates of its model SBCs (Fig. 3A, top row). Without synaptic depression, the model MSO neuron responded strongly to its best-ITD[2] (zero ITD) across the AM cycle: most strongly at peak sound-energy, slightly less strongly during rising energy, and less strongly still during falling energy[3] (Fig. 3A, lower three rows; in a 3-way comparison of spike counts for best-ITD during rising vs. peak vs. falling energy, $p < 10^{-9}$ at each modulation frequency). The slightly larger suppression in responses to best-ITD during falling energy is supported by adaptation in the AN model, and K_{LT}-based refractory effects in the model SBCs and MSO neuron.

Next, we introduced synaptic depression in supra-threshold synapses of model SBCs, producing adaption in their spike rates (Fig. 3B, top row). This adaptation followed a similar time-course across modulation rates, such that at low-to-moderate rates (4 to 16 Hz) most of the adaptation occurred by the peak in the AM cycle, and at moderate-to-high rates (16 to 64 Hz) gradual adaptation is evident across much of the AM cycle. Consistent with physiological and behavioural data at 600 Hz (Dietz et al., 2014; Hu et al., 2017), the model MSO neuron responded to its best-ITD strongly during rising amplitude, weakly to moderately (and broadly) at peak energy, and only weakly during falling energy (Fig. 3B, lower three rows; in this 3-way comparison of spike counts, $p \leq 0.003$ at each modulation frequency). Adapting similarly to the model SBCs, the model MSO neuron had higher spike counts for best-ITD during rising energy vs. peak energy at 4- to 16-Hz modulation ($p < 0.02$), and during peak vs. falling energy at 16- to 64-Hz modulation ($p < 0.0025$).

---

[1] Characteristic frequency (CF) is the frequency at which a neuron has its lowest threshold in sound pressure level for an action potential; CF is inherited from the frequency-tuning of the inner ear, and the resulting CFs of auditory nerve fibres that distally drive a neuron.

[2] The best-ITD of a neuron is the ITD of highest spike rate, typically measured with unmodulated stimuli of equal amplitude at the two ears, such as the unmodulated binaural beats in Dietz et al. (2014). The nonlinear model MSO neurons of this study have best-ITDs equal to zero.

[3] AMBB stimuli had increasing IPD (positive beat direction): “best-ITD during rise” denotes *start-IPD* = 270°; “best-ITD at peak” denotes *start-IPD* = 180°; and “best-ITD during fall” denotes *start-IPD* = 90°.

Thus, consistent with behavioural data at the same, 600-Hz, sound frequency, a model that includes synaptic depression between the auditory nerve and cochlear nuclei that drove bilateral inputs to the MSO neuron, responds preferentially to ITDs in the rising energy of sounds. Without this synaptic depression, the preference for spatial cues in rising sound-energy is absent, suggesting a critical role for monaural input pathways to binaural neurons in emphasising spatial information that is likely to be reliable in reverberant listening conditions.

Sensitivity to ITDs conveyed in the temporal fine structure (TFS) of sounds is subject to highly specialized biophysical constraints, constraints that are generally thought to hold for sound frequencies up to 1400 Hz (higher in some non-human species), covering the range over which ITDs in the TFS are discriminable. Therefore, given the emphasis to ITDs contained in the rising energy of modulated sounds at 600 Hz—concordant with perceptual and brain-imaging data obtained at a similar frequency (Dietz et al., 2013)—we might expect a similar emphasis at the lower sound-frequency of 200 Hz. Nevertheless, despite our model implementing identical biophysical elements, we observed a very different pattern of results at 200 Hz compared to 600 Hz. These differences are reminiscent of behavioural data (Hu et al., 2017): at 600 Hz, human ITD-sensitivity is strongest during rising energy with progressively less ITD sensitivity for peak and falling energy; but at 200 Hz, ITD sensitivity is strong for both peak and rising energy, and again weak for falling energy.

Employing the same model with monaural depressing synapses, identical to those which emphasised ITDs in the rising energy at 600 Hz, we assessed the relative emphasis of ITD cues in the modulation cycle of a 200-Hz AMBB stimulus for the modulation rates of 8 and 20 Hz employed by Hu et al. (2017). At 200 Hz, the model SBCs showed slight adaptation in spike rate across the AM cycle (Fig. 4A, top row), less adaptation than at 600 Hz. Spike counts in the model MSO neuron (Fig. 4A, lower three rows) matched the patterns of significance in human ITD-detection for AM stimuli at 200 Hz (Hu et al., 2017): spike counts were slightly higher but not significantly for best-ITD during peak vs. rising energy, and spike counts were significantly different (higher) only for best-ITD at peak energy vs. falling energy ($p = 0.0079$ for modulation at 8 Hz; $p = 0.0014$ for modulation at 20 Hz).

Noting recent evidence of heterogeneity in membrane speeds in the very low-frequency region of MSO (Bondy & Golding, 2018), we applied AMBBs at 200 Hz to our nonlinear brainstem model, now modified with a moderate-speed model MSO neuron having slower dendritic and somatic membranes, derived by dividing membrane conductance values by three. The identical model SBCs again adapted less (Fig. 4B, top row) than at 600 Hz. Synaptic strength in this moderate-speed model MSO neuron was divided by two for low spike rates for best-ITD during falling energy, while retaining moderate spike rates for best-ITD during rising and peak energy (Fig. 4B, lower three rows). The moderate-speed MSO neuron then produced very few spikes at ITDs distant from best-ITD, yielding spike counts for best-ITD at peak energy that were equal to, or slightly less

than, spike counts for best-ITD during rising energy. The moderate-speed model MSO neuron slightly augmented patterns of significance in human ITD-detection (Hu et al., 2017): spike counts were not significantly different for best-ITD during rising vs. peak energy, and were significantly lower for best-ITD at falling energy (8-Hz modulation: rise vs. fall, $p = 4.08 \times 10^{-05}$; peak vs. fall, $p = 0.0049$; 20-Hz modulation: rise vs. fall, $p = 1.53 \times 10^{-09}$; peak vs. fall, $p = 1.53 \times 10^{-09}$).

Our nonlinear brainstem model accounts for the emphasis of rising energy in ITD encoding by MSO neurons at 600 Hz, and is consistent with the shift from very strong human ITD-sensitivity during rising energy at 600 Hz, to strong sensitivity during both peak and rising energy at 200 Hz, and with weaker sensitivity during falling energy at both frequencies. As the stimulus frequency is reduced from 600 to 200 Hz, the reduction in suppression at the peak of the AM cycle is likely due to decreased activation of K_{LT} channels at the lower frequency, and the inherently longer quiet periods between intense input spiking (Chung et al., 2015; Colburn et al., 2009).

A hemispheric-difference model incorporating synaptic depression correctly lateralises amplitude-modulated stimuli with temporally specific interaural time differences

To test whether synaptic depression in combination with heterogeneous neuronal properties is consistent with the behavioural observations of Hu et al. (2017), we employed a neural spiking model of both brain hemispheres, incorporating linear model neurons in both cochlear nuclei and MSO stages (Remme et al., 2014), to generate a neuro-metric measure of lateralisation. Following the model concept of Dietz et al. (2009), two hemispheric channels ('left' and 'right') comprised distinct MSO model neuron populations (50 neurons each) whose inputs from 'contralateral' cochlear nucleus were delayed by 0.125 cycles of interaural phase difference (IPD, equal to ITD x frequency) (McAlpine et al., 2001) such that they spiked preferentially for sounds arriving from contralateral spatial locations (Fig. 5). The difference in spike rate between hemispheres, calculated (in 5-ms epochs) as a signed d', was then used as a neuro-metric measure of lateralisation (see Methods) (Devore et al., 2009; McAlpine et al., 2001). A uniform synaptic depression (Rudnicki & Hemmert, 2017; $u = 0.55$) was implemented independently at each synapse between an AN fibre and its target SBC of the cochlear nucleus. Sinusoidally amplitude-modulated (SAM) tones, with non-zero, positive ITDs positioned within the rising, peak, or falling phases of the AM cycle [as per Hu et al., (2017), and see above], were presented to the linear model (Fig. 6). As in the behavioural study, the tone frequency (200 Hz or 600 Hz) and the AM rate (8 Hz or 20 Hz) were varied, as was the proportion (20% or 40%) of the AM cycle containing a non-zero ITD.

At 600 Hz, MSO neurons with fast resonant properties (i.e. akin to the nonlinear model with 100% K_{LT} conductance) demonstrated a strong sensitivity to the onset ITD, as indicated by the increased number of instances when the AM stimulus containing a (right-leading) +150 µs ITD in its rising-energy phase was correctly lateralised [d' > 1 in any bin was considered a correct lateralisation of a signal from the 'right' (Fig. 6 *pink*

vertical bands)]. This ITD sensitivity decreased across the AM cycle reaching a minimum when non-zero ITDs were restricted to the falling-energy phase (Fig. 6 *bottom row*). The trend for onset dominance at 600 Hz was true for both modulation rates (8 Hz and 20 Hz) and when the proportion of the AM cycle containing non-zero ITDs was either 20% or 40% (Fig. 7 *right column*). Notably, in all three portions of the AM cycle, the number of correct lateralisations decreased as the proportion of an AM cycle containing a non-zero ITD was decreased from 40% to 20% (Fig. 7 *right column*, blue).

When an amplitude-modulated 200-Hz tone (with a right-leading, non-zero ITD of +300 µs) was presented to the fast MSO model, the onset dominance observed at 600 Hz was replaced by an increased weighting of ITD cues towards the peak of the AM cycle (Fig. 6). Indeed, the frequency of correct lateralisations at the peak of AM cycles was either equal to (8-Hz AM stimuli) or higher (20-Hz AM stimuli) than that observed at corresponding onset phases (Fig. 7 *middle column*). As at 600 Hz, reducing the proportion of the AM cycle containing a non-zero ITD from 40% to 20% also generated fewer correct lateralisations across all AM phases (Fig. 7 *middle column*, blue). Whilst the fast-resonant MSO model could reproduce behavioural trends observed by Hu *et al*. (2017) at both carrier frequencies, it generated fewer correct lateralisations overall at 200 Hz (Fig. 7 *middle* and *right columns*). We therefore presented the same 200-Hz carrier stimulus to the linear model MSO neurons with slower, less resonant membrane properties (Fig. 6). This generated a higher number of correct lateralisations overall, while maintaining, or even augmenting, the maximum weighting of ITD at the peak AM energy in all conditions tested (Fig. 7 *left column*).

Our linear model, therefore, suggests that a slower, more integrative MSO neuron may assist lateralisation in lower-frequency channels, in particular by extending the extraction of ITD information towards peak energy in the AM cycle, where human ITD detection is best when comparing at 200 Hz across the AM cycle (Hu et al., 2017). Additionally, our model qualitatively matched the behavioural data by demonstrating a higher propensity to lateralise stimuli of either frequency when 40% vs. 20% of the AM cycle contained a non-zero ITD.

Synaptic depression improves the hemispheric-difference model's ability to lateralise speech signals in artificially reverberant conditions

If neural adaptation prior to binaural integration contributes to accurate source localisation, we would expect it to aid lateralisation of speech in a reverberant room, by emphasising ITD cues associated with direct sound over reverberation (Fig. 1 *far left*) (Dietz et al., 2014). Speech signals were presented to the neural model of lateralisation under simulated anechoic (non-reverberant) and reverberant conditions. We then observed whether the model correctly lateralised the side to which the source was located in the presence or absence of synaptic depression in SBCs of the cochlear nuclei.

In the anechoic condition (Fig. 8A), a word token ('church') was presented to the non-adapting linear model from a virtual location 1 metre away from the target and 30 degrees to the right of the midline. The speech waveform—gammatone-filtered at 600 Hz (Fig. 8A *top row*)—was more intense at the right ear (red), and led the left-ear signal (blue) by an average ITD of +363 µs (Fig. 8A *top row*, dark grey). This generated a larger response from model neurons in the 'left' MSO (Fig. 8A *middle row*) and multiple bins throughout the speech token where d' values were > 1 (Fig. 8A *pink vertical bands*), indicating that the talker was correctly lateralised to the right in the non-adapting, linear model.

In the reverberant condition, the same word token was presented to the non-adapting linear model from the same angle and distance, however it was also followed by two delayed copies generated by early reflections off virtual walls added behind, and to the left of, the listener (Fig. 1B *far left*). This mixture of sound arriving direct from the source and early reflective copies led to a more complex waveform (including a more intense signal at the 'left ear' (DRR: -4dB; mean ILD: -3 dB)) whose running ITDs fluctuated between extremely large positive and negative values, culminating in an average ITD of +711 µs, a value outside the human physiological range. Although the resulting d' output exceeded 1 in eleven epochs coinciding with waveform onsets (Figure 8A, *pink vertical bands*), it was also briefly lower than -1 in the opposite direction during three intermediate epochs (Fig. 8A, *blue vertical bands*), indicating a potentially ambiguous left-right localisation output.

We then increased adaptation by introducing synaptic depression to the monaurally driven inputs from AN fibres to the cochlear nuclei of the linear model, and presented the same reverberant stimulus (Fig. 8B) as in the previous condition. Despite the potentially confounding ITD (and ILD) cues (Fig. 8B *top row*), adding adaptation in the monaural input pathways enhanced the performance of the binaural model MSO neurons, and generated correct lateralisations of the true source to the right: the d' output exceeding 1 in six time bins without ever crossing below -1 (Fig. 8B *bottom row*). When the number of incorrect lateralisations was tallied over a hundred presentations of the reverberant stimulus (Fig. 8C), incorrect lateralisations were observed five times less frequently when synaptically depressing AN inputs were included in the linear model (mean incorrectly lateralised time for non-depressing AN inputs = 6.75 ± 0.50 ms/presentation vs. mean incorrectly lateralised time for depressing AN inputs = 1.3 ± 0.24 ms/presentation). This neuro-metric profile agrees with the psychoacoustic percept as has been previously described for similar word tokens (Dietz *et al*., 2014), suggesting that synaptic adaptation helps to lateralise ethological sounds in reverberant environments.

DISCUSSION

Frequency-dependent emphasis of spatial cues in reverberant environments

The ability to determine the source of a sound in reverberant environments, where reflections from acoustically opaque surfaces can degrade the accuracy of auditory spatial cues, is critical to prey and predator alike, and an important facet of effective communication, including in human listeners. Here, employing physiologically realistic models of brainstem auditory neurons, we find that accurate localisation depends on both synaptic depression in neurons of the cochlear nucleus that project bilaterally to the medial superior olive (MSO, the site of primary binaural integration)—as well as on K_{LT} channels in the cochlear nuclei and the MSO that activate more strongly at relatively higher sound-frequencies. Combined, these suppressive mechanisms underpin the ability of MSO neurons to respond preferentially to ITDs conveyed in the early, rising energy of low-frequency sounds, especially for the 'dominant' frequency region for ITD encoding spanning 500-1200 Hz, a range high in reverberation (Traer & McDermott, 2016) but which also generates the lowest behavioural thresholds for ITD discrimination (Brughera et al., 2013; Klumpp & Eady, 1956).

Specifically, synaptic depression at monaurally driven model spherical bushy cells (SBCs) of the cochlear nucleus, and K_{LT} channels in these, and in binaurally driven model MSO neurons, accounted for the physiologically measured emphasis of ITDs conveyed during rising sound energy for frequencies near 600 Hz. Both linear and nonlinear brainstem models generated the observed readout-window for ITD cues during each cycle of amplitude-modulated sounds centred near 600 Hz (Dietz et al., 2013, 2014); these models successfully glimpsed the ITD cues conveyed in the temporal fine structure during the rising portion of the AM cycle, suppressing spatial information conveyed in all later portions of the AM cycle, including the portion with highest energy.

Our models are consistent with the frequency dependence observed behaviourally when ITDs are applied to different segments of the AM cycle (Hu et al., 2017). At 600 Hz, spike counts in model MSO neurons reflected the strong sensitivity by human listeners for ITD during rising sound-energy. Reflecting the strong human ITD-sensitivity during peak and rising sound-energy at 200 Hz, as well as reduced suppression by K_{LT} channels at a lower sound frequency, model spike counts were strong during peak and rising energy at 200 Hz, and were suppressed at both frequencies during falling sound-energy when reverberation is strongest compared with direct sound. Intrinsic K_{LT}-channel properties, that decrease K_{LT} activation more during the longer quiet periods between periodic bursts of intense input spiking at lower stimulus frequencies (Chung et al., 2015; Colburn et al., 2009), are likely contributing mechanisms for the shift from very strong human ITD-sensitivity during rising energy at 600 Hz, to strong sensitivity during both peak and rising energy at 200 Hz, with weaker sensitivity during falling energy at both frequencies.

The K_{LT} activity supports ITD sensitivity (Mathews et al., 2010), and combines with synaptic depression to transmit reliable, and suppress unreliable, auditory-spatial information, accounting for a natural frequency-profile in reverberant energy. As sound-

frequency decreases below 1500 Hz, and reverberation times decrease to a minimum near 200 Hz for many natural environments (Traer & McDermott, 2016), suppressive K_{LT} activity also decreases (Chung et al., 2015; Colburn et al., 2009); and to the extent that spike rates of AN fibres decrease with lower sound-frequency, synaptic depression in neurons of the cochlear nucleus may also decrease. At very low frequencies, including 200 Hz, where reverberant energy in natural, outdoor scenes is low, suppression of spatial information during peak energy is less important. However, in modern, indoor listening environments, characterised by enclosures with highly reflective walls, reverberation can be high even at very low frequencies. Indoor reverberation combined with weakly suppressing neurons, and with increasing perceptual thresholds for ITD at sound frequencies below 500 Hz (Brughera et al., 2013), may contribute to the observed low perceptual weighting of very low frequencies when localizing broadband sound (Ihlefeld & Shinn-Cunningham, 2011).

Consistent with recently reported heterogeneity in temporal properties of neurons in the low-frequency region of MSO (Bondy & Golding, 2018), both moderate and fast nonlinear model MSO neurons reflected human ITD sensitivity at 200 Hz (Hu et al., 2017). The moderate-speed model neuron with its lower membrane conductance (in dendrites and soma, including K_{LT} conductance) is somewhat slow compared with early measurements of MSO neurons (Scott et al., 2007), but of moderate speed compared with recent measurements (Mikiel-Hunter et al., 2016; Remme et al., 2014). Reduced, but still viable, K_{LT} activity provided ITD sensitivity at 200 Hz in the nonlinear model, while readily combining with synaptic depression in model neurons of the cochlear nucleus to suppress responses to spatial information during falling energy, late in the AM cycle. Introducing slower membrane properties to the linear model MSO neurons boosted the hemispheric-difference model's ability to lateralise 200-Hz signals, especially at their energetic peak, suggesting a functional benefit to a reduced K_{LT}-density at lower frequencies. The relatively slow membrane time-constants of many MSO neurons tuned to very low sound-frequencies are still rapid compared to other neuron types; and by not expressing more ion-channels than required, MSO neurons encode ITD efficiently, providing a cellular energetic advantage.

Synaptic adaption enhances correct lateralisation of reverberant speech

Early reflections—those following the direct signal by 50 ms or less (Bradley et al., 2003)—can disrupt low-frequency sound localisation in realistic auditory environments by conveying spurious ITD cues in the TFS of sounds (Gourévitch & Brette, 2012). Despite this, normal-hearing listeners are highly adept at locating the true source of a sound in reverberant environments including speech in reverberant environments. By demonstrating that the addition of synaptic adaption to a hemispheric 2-channel model emphasises ITD cues in sound onsets, our data suggest one means by which complex sounds, including speech, are reliably lateralised.

Whilst our adapting hemispheric-difference model correctly lateralised reverberant speech, this was based on its d' output surpassing a threshold value in only six of the

ninety 5-ms bins where a correct ITD cue was available. Nevertheless, although this neuro-metric measure may not appear particularly robust, it should be remembered that only the 600-Hz frequency channel was examined for the lateralisation of speech. Studies of the precedent effect—the suppression of auditory information, including spatial information—for late-arriving sounds—and localisation in reverberant conditions have demonstrated strong weighting of cues in the 500-750 Hz frequency range (Ihlefeld and Shinn-Cunningham, 2011; Shinn-Cunningham et al., 1995; Xia et al., 2010), but they also highlight the inability of subjects to localize pure tones in echoic conditions, suggesting that enhanced localisation of speech in reverberant conditions involves localisation cues across the low-frequency spectrum. Adding a midbrain stage where neurons in the inferior colliculus (IC), for example, receive inputs from MSO neurons across a wide range of CFs may indicate the extent to which synaptic adaptation in the cochlear nucleus improves the localisation of speech in reverberant environments.

Finally, we note that onset-cue dominance is demonstrated extensively in the precedence effect (Wallach et al., 1949), where suppression of neural responses to a brief lagging stimulus occurs over a range of delays between leading and lagging stimuli. A future localisation model combining monaural synaptic depression, potentially with inhibitory echo-suppressive mechanisms, such as delayed inhibition to the midbrain (Burger & Pollak, 2001; Kidd & Kelly, 1996; Pecka et al., 2007), and feedforward inhibition in the brainstem from dorsal to ventral cochlear nucleus (Keine et al., 2017) and from ventral cochlear nucleus to MSO via the trapezoid body (Adams & Mugnaini, 1990; Cant & Hyson, 1992; Smith et al., 1991), provides one means of exploring whether these mechanisms act cooperatively in robust onset-ITD processing in reverberant conditions.

METHODS

Nonlinear brainstem model

This computational model (Fig. 2B) incorporates an MSO neuron and its excitatory input pathways including synaptic depression in the VCN. Amplitude-modulated binaural beats (AMBBs) were applied as acoustic stimuli to an auditory periphery model (Zilany et al., 2014) implemented in Matlab (www.mathworks.com), including 12 left and 12 right model AN-fibres that incorporate power-law adaptation, and project depressing synaptic inputs to model SBCs of the VCN (Rothman & Manis, 2003c; Rudnicki & Hemmert, 2017). Three model AN fibres drive each model SBC. Four left and four right SBCs project non-depressing, excitatory synaptic inputs to a multi-compartment model MSO neuron (Brughera et al., 2013) with recent updates. These model brainstem neurons are implemented as Hodgkin and Huxley (1952) models in the Python 3.7 Brian2 Neural Simulator (Stimberg et al., 2013).

Model SBCs with synaptic depression in the VCN reflect physiological constraints

SBCs are modelled as point neurons (Rothman & Manis, 2003c, Type II compartment, adjusted for temperature 37°C) with three independent, excitatory synapses each driven by a model AN fibre. Each model synapse is an excitatory synaptic conductance in series with an excitatory reversal potential of 0 mV. Each input spike causes a variable increment in its synapse's excitatory conductance which decays exponentially with time constant 0.2 ms. The increment (i.e., the synaptic strength) is variable due to synaptic depression. Non-depressed synaptic strength is given in Table 2. Synaptic depression at the individual synapses is modelled as in Rudnicki and Hemmert (2017): immediately after an input spike and its associated increment in excitatory conductance, the synaptic strength is multiplied by $1 - u$, where $u = 0.5$, comparable to measurements *in vitro* ($u = 0.6$ in Oleskevich et al., 2000; $u = 0.4$ in Wang & Manis, 2008; Yang & Xu-Friedman, 2009). This synaptic depression recovers with a single-exponential function with time constant 25 ms, equal to the fast component of recovery measured in Wang and Manis (2008) which accounted for 70 percent of synaptic depression. In their measured Endbulb of Held synapses, the remaining 30 percent of synaptic strength recovered with a time-constant near 2 s, and would remain almost completely depressed during our modelled responses to ongoing AMBB stimuli.

Modelled synaptic depression contributed to realistic responses to AMBB stimuli, while providing consistency with phase-locking to AM in the auditory nerve (Joris & Yin, 1992) which is then maintained by primary-like neurons (including bushy cells) in the cochlear nucleus (Rhode & Greenberg, 1994). Three inputs per model SBC (compared with a single input) increased spike rates, a factor in matching the low range of synchrony to AM, and in supporting a realistically small number of excitatory inputs (Couchman et al., 2010) to the model MSO neuron.

Model MSO neuron

The model MSO neuron has separate compartments representing the bilateral dendrites, soma, and axon of the neuron (Zhou et al., 2005). As neuronal delays and inhibition are not implemented, the soma is simplified, being bilaterally symmetric with a spherical, iso-potential soma. Best-ITD and best-IPD are zero. Parameter values for each compartment (Table 1) are explained below. Certain parameters (Table 2) were adjusted for the four conditions yielding results shown in Figs. 3-4.

Eight excitatory synapses are located on the model MSO bilateral dendritic membranes, one each at 42.5, 47.5, 52.5, and 57.5% of the dendritic length. Each synapse is modelled as a variable conductance (g_E) in series with an excitatory reversal potential (V_E) of 0 mV. The synapses are non-depressing, and each is driven by a single model SBC from the same side. An input spike increments its synaptic conductance by a fixed amount, and the conductance then decays exponentially with time constant 0.4 ms. Synaptic strength (Δg_E), which is the fixed conductance-increment, was adjusted according to CF, the overall membrane conductance including g_{KLT}, and the presence or absence of synaptic depression in the cochlear nucleus.

Within the model MSO neuron, its functional electric-circuit unit the sub-compartment (Fig. 2C) connects to its neighbouring sub-compartments via series resistivity (R_S) representing the neuron's internal cytoplasm. Within each sub-compartment, the neural membrane is modelled as a transmembrane capacitance (C_M) in parallel with ion-channel populations: sodium (Na^+) channels, hyperpolarization-activated cyclic nucleotide (HCN, I_h) channels for non-specific cations, and K^+_{LT} (K_{V1}) channels. Each ion-channel population is represented as a reversal potential (V_{Na}, V_H, V_{KLT}) in series with a conductance (g_{Na}, g_H, g_{KLT}). With the exception of a fixed leakage conductance (g_L) representing voltage-insensitive ion channels, each conductance value is equal to a maximum conductance multiplied by voltage-sensitive activation and inactivation gating variables with integer exponents (Hodgkin & Huxley, 1952).

Kirchhoff's current equation, applied to the model MSO neuron, states that the sum of the membrane and series currents exiting any point in the model neuron is zero:

$$0 = i_{C_M} + i_{Na} + i_{KLT} + i_H + i_L + i_{S1} + i_{S2}$$

Series currents, i_{S1} and i_{S2}, are calculated by the Brian2 simulator according to the voltage differences, resistivity, and geometry of the related sub-compartments.

Capacitive membrane current (i_{C_M}) increases with membrane capacitance and the time-derivative of membrane potential V:

$i_{C_M} = C_M \frac{dV}{dt}$.

Leakage current: $i_L = g_L(V - V_L)$.

Sodium current (Scott et al., 2010): $i_{Na} = g_{Na}(V - V_{Na})$.

Na^+ conductance: $g_{Na} = gmax_{Na} m^4 (0.993h + 0.007)$.

Each voltage-sensitive ionic conductance has a constant maximum conductance value ($gmax$) (Table 1), and has voltage and time dependencies defined by a subset of the activation/inactivation variables m, h, w, z, and r, each with a rate of change governed by a first-order differential equation with a time constant divided by the Q_{10} temperature factor of $3^{(T-22)/10}$, where T is set equal to human body temperature, 37°C. For sodium:

$$\frac{dm}{dt} = Q_{10}(m_\infty - m)/\tau_m$$

$$\frac{dh}{dt} = Q_{10}(h_\infty - h)/\tau_h$$

For model Na^+ channels, steady-state activation (m_∞) and inactivation (h_∞), and their respective time constants (τ_m and τ_h) in millisceconds, are functions of membrane potential in millivolts, equal to those in Scott et al. (2010):

$$m_\infty = 1./(1 + \exp((V + 46.) / (-11.)))$$

$$\tau_m = ((0.141 + (-0.0826 / (1 + \exp((-20.5 - V) / 10.8)))) / 3.)$$

$$h_\infty = 1./(1 + \exp((V + 62.5) / 7.77))$$

$$\tau_h = ((4. + (-3.74 / (1 + \exp((-40.6 - V) / 5.05)))) / 3.)$$

Low-threshold potassium current (Mathews et al., 2010):

$$i_{KLT} = g_{KLT}(V - V_{KLT}).$$

$$g_{KLT} = gmax_{KLT} w^4 z.$$

$$\frac{dw}{dt} = Q_{10}(w_\infty - w/\tau_w$$

$$\frac{dz}{dt} = Q_{10}(z_\infty - z)/\tau_z$$

$$w_\infty = 1./(1 + \exp((V + 57.34)/(-11.7)))$$

$$\tau_w = ((21.5/((6.* \exp((V + 60.)/7.)) + (24.* \exp(-1 * (V + 60.)/50.6)))) + 0.35)$$

$$k = 0.27$$

$$z_\infty = ((1. - k)/(1 + \exp((\mathrm{V} + 67.)/6.16))) + k$$

$$\tau_z = ((170./(5.* \exp((V + 60.)/10.) + \exp((V + 70.)/8.))) + 10.7)$$

Non-inactivating HCN current (Rothman & Manis, 2003c):

$$i_H = g_H(V - V_H).$$

$$g_H = gmax_H r \ .$$

$$\frac{dr}{dt} = Q_{10}(r_\infty - r)/\tau_r$$

$$r_\infty = 1./\ (1 + \exp((\mathrm{V} + 76.)\ /\ 7.))$$

$$\tau_r = ((100000./\ (237.* \exp((\mathrm{V} + 60.)\ /\ 12.) + 17.* \exp(-(\mathrm{V} + 60.)\ /\ 14.))) + 25.)\ .$$

Approach for Setting Maximum Conductance Values of MSO Ion-Channels

By selecting equilibrium membrane potentials (V_{REST}) and equilibrium ratios of i_{Na} to i_H to i_{KLT} in each compartment (0:1:-1 in dendrites; 1:4:-5 in soma, 1:1:-2 in axon), and knowing C_M and membrane time constants, the current equation yielded a maximum conductance value for each ion-channel in each compartment that produced: K^+_{LT} activation in dendrites and soma; Na^+ inactivation in soma; Na^+ not inactivated in axon.

Model MSO membrane time constants

In the resulting model MSO neuron of this study, with bias current to centre the somatic membrane potential at -57.6 mV, slightly above V_{REST}, the somatic membrane impedance magnitude as a function of frequency was lowpass with a resonance at 480 Hz. Although this resonance frequency is slightly higher than typical for MSO neurons (Mikiel-Hunter et al., 2016; Remme et al., 2014), this 480-Hz resonance frequency multiplied by 2π converts to an angular frequency of 3016 radians/s, with its reciprocal yielding a membrane time constant of 0.33 ms that is within the range of MSO neurons (Scott et al., 2007). For AMBB stimulation centred at 200 Hz, to reflect the range of MSO membrane speeds at very low CF (Bondy & Golding, 2018), slower somatic and dendritic membranes were also modelled by dividing the fast membrane conductances by three to yield a somatic membrane time constant of approximately 1 ms.

AMBB period histograms

For each AMBB period histogram from the nonlinear model, spikes were counted in forty non-overlapping bins, each covering 1/40 of the AM cycle. Histograms showing results from the nonlinear model were not smoothed. Spike rates were calculated by dividing the spike count by the total time duration for each bin, across the multiple periods of eight different stimulus presentations of 0.75-second duration.

Eight starting-phases at decrements of 45° efficiently implemented IPDs

At each carrier and modulation-frequency combination, 8 acoustic stimuli, each with 1 of 8 carrier *starting-phases* (0°, -45°, -90°, …, -315°), were applied at each model ear, AN-fibre, and SBC. For computational efficiency, *start-IPD* (the IPD at zero amplitude) was achieved by pairing spike times from each *starting-phase* in one ear, with the spike

times from the other ear having the appropriate difference in *starting-phase*. While maintaining the proper *start-IPD*, each AMBB period histogram pooled spikes resulting from stimuli using the 8 *starting-phases* spanning each carrier period. Thus while the histograms show the effects of *start-IPD*, they do not show phase-locking to fine structure of the carriers.

In this study, "best-ITD during rise" denotes *start-IPD* = 270°; "best-ITD at peak" denotes *start-IPD* = 180°; and "best-ITD during fall" denotes *start-IPD* = 90°.

Chi-squared tests for statistically significant differences in model spike counts

For each condition and modulation frequency, a chi-squared test for one and two degrees of freedom (2-way and 3-way comparisons) compared spike counts from the nonlinear model MSO neuron stimulated with AMBBs with best-ITD occurring during rising vs. peak vs. falling amplitude. Each chi-squared test yielded a probability (*p*) of the null hypothesis that the differences in spike count occurred randomly.

Synchrony index and Rayleigh statistic for significance

For each modulation frequency and *starting-phase* combination, the synchrony index (*SI, R*) of spike times (t_i) (Johnson, 1980) with respect to their phase θ_i within the AM cycle (with modulation frequency f_m and modulation period $T_m = 1/f_m$) was calculated:

$$\theta_i \triangleq 2\pi f_m mod(t_i, T_m)$$

Synchrony index $= \sqrt{(\sum_{i=1}^{N} \cos\theta_i)^2 + (\sum_{i=1}^{N} \sin\theta_i)^2}$, where *N* is the spike count.

The Rayleigh significance statistic ($2NR^2$) was then calculated and converted to a *p* value, the probability of the null hypothesis that the AM period-histogram of spike times resulted from a uniform distribution (Rhode, 1976):

p	0.10	0.05	0.025	0.01	0.001
$2NR^2$	4.605	5.991	7.378	9.210	13.816

A model of lateralisation using linear-membrane models

To study how synaptic adaptation and biophysical profiles could impact lateralisation of acoustic stimuli in anechoic and reverberant conditions, we constructed a model of lateralisation whose brainstem stages (SBCs and MSO neurons) are represented by point neurons with linear biophysical properties (Remme et al., 2014). Two stimuli were presented to the model: the first, a SAM tone with non-zero ITDs inserted at different phases of the AM cycle, has been used behaviourally to determine how carrier frequency affects ITD sensitivity over the time course of an AM cycle (Hu et al., 2017); the second, a natural speech stimulus with early reflections, has been proposed as a simple simulation of a reverberant environment where correct lateralisation is only

possible if ITD cues in the onset waveform are processed preferentially (Dietz et al., 2014).

SAM tones with non-zero ITDs inserted at different phases of AM cycle

This stimulus was generated according to the Hu *et al.* study (2017) that first used it. In summary, SAM tones were created from pulse trains that had been bandpass-filtered and then amplitude-modulated. Non-stationary ITDs were added by temporally shifting individual pulses in one of either the rising, peak or falling phases of the AM cycle. For the purposes of this study, non-stationary ITDs of +300 μs and +150 μs were inserted into SAM tones of either 200-Hz or 600-Hz carrier frequency respectively. These ITD values were chosen based on individual behavioural thresholds in the Hu *et al.* study (2017). At each carrier frequency, four conditions were tested: two different AM frequencies (8 Hz or 20 Hz) and two different proportions of rising, peak or falling AM phases with a non-zero ITD inserted (20% or 40%). These parameters match those tested behaviourally by (Hu et al., 2017).

A natural speech stimulus with early reflections

A single monosyllabic consonant-vowel nucleus-consonant (CNC) word token, ‘church’, (spoken by an Australian-English, female voice), was selected as a stimulus (Fig. 1 and Fig. 8). To simulate its arrival from a location one metre away to the right of the midline in an anechoic environment, the word’s waveform (normalized at 70dB) was convolved with a large pinnae HRTF for +30° azimuthal location (CIPIC database, UC Davis). A reverberant space was simulated by adding to the direct (+30 deg) pathway two early reflections off virtual walls stood behind and to the right of the subject (Fig. 1). These reflected copies of the words arrived from -60° and -130° angles (again convolved with large-pinnae HRTFs from the CIPIC database) and were delayed by 4ms and 8ms respectively to reproduce their elongated paths (Dietz et al., 2014).

Linear model circuitry

The linear model, like its nonlinear counterpart, consists of three stages: an initial auditory periphery stage (the middle/inner ear and AN fibres) and two brainstem stages (VCN and finally the MSO). Unlike the nonlinear model however, the linear model incorporates two parallel circuits, representing the ‘left’ and ‘right’ MSO nuclei and their inputs. The difference in output rate between these parallel circuits is utilised to calculate lateralisation of the natural stimulus. The only discernible difference between ‘left’ and ‘right’ circuits is the contralateral $\frac{1}{8}$ pi shifts in their MSO neurons’ IPD tuning curve peaks which are generated by delaying the outputs from the auditory periphery stage ipsilateral to the MSO nucleus in question. This shift is considered physiological given experimentally observed binaural tuning curves from the auditory brainstem whose optimal characteristics have been corroborated theoretically (McAlpine et al., 2001).

Auditory periphery stage

Peripheral processing of the acoustic stimuli in the linear model is also performed using an implementation of the Zilany *et al.* (2014) cat middle/inner ear model in Matlab. All parameters in the periphery are identical between the nonlinear and linear models. Outputs from the auditory periphery are restricted to 200-Hz/ 600-Hz for the pulse-shifted SAM tones (matching their SAM tone's carrier frequency) and 600-Hz for lateralisation of the speech stimulus with early reflections. This latter value was chosen based on it laying in the middle of the 500-750 Hz frequency range in which TFS-ITD sensitivity is considered strongest (Ihlefeld & Shinn-Cunningham, 2011). When exploring the effects of AN adaptation on the lateralisation of a speech stimulus, the same auditory periphery simulations are used for both the adapting and non-adapting inputs to the CN stage.

Linear-membrane model properties

Single-compartment, linear-membrane models are used to represent both SBC and MSO neurons in the linear model. These linear-membrane models can incorporate two dynamic currents as previously described in (Remme et al., 2014): I_w, a resonant current and/or, I_n, an amplifying current. The current balance equation for the general neural model is as follows:

$$\mathrm{c}\frac{dv}{dt} = -g_{\mathrm{M}}v - g_{\mathrm{w}}w_{\mathrm{w}} + g_{\mathrm{n}}w_{\mathrm{n}} + \mathrm{I}_{app}(t)$$

Where c is capacitance (picofarads), g_M is total membrane conductance (nanosiemens) and I_{app} is the current (picoampere) applied to the neural model. The resonant and amplifying currents, I_w and I_n, are described by their conductances, g_W and g_n (in nanosiemens); the dynamics of their gating variables, w_w and w_n (in millivolts) are described generically by the equation:

$$\tau_{\mathrm{w}}\frac{dw_{\mathrm{w}}}{dt} = v - w_w$$

$$\tau_{\mathrm{w}}\frac{dw_{\mathrm{w}}}{dt} = v - w_w$$

Where w_x is the gating variable whose associated time constant is τ_x (milliseconds).

Two parameter sets are chosen to represent the apparent bimodality of subthreshold biophysical properties in the gerbil MSO *in vitro* (Bondy & Golding, 2018). These example models can be broadly characterized by their impedance profiles as either band-pass with a high resonant-frequency or band-pass with a low resonant-frequency (Remme et al., 2014). While the band-pass model with a high-resonant frequency is

exclusively chosen to represent SBCs, both model types are used to represent MSO neurons.

Synaptic properties and adaptation

As in the nonlinear model, 3 independently-simulated AN-fibres provide excitatory drive to each SBC in the linear model. Excitatory unitary conductances at this synapse are modelled as alpha functions with an exponential time course of 0.2ms. Synaptic depression with a single exponential recovery is also implemented at this stage, matching the adaptation observed in the nonlinear model functionally thanks to its similar parameters (u = 0.55; τ = 25 ms).

Each MSO neuron receives 4 (independently-simulated) excitatory VCN inputs from "either ear" in the linear model. Alpha functions are again used to model the unitary excitatory conductances at this synapse with time courses selected relative to the resonant class of the MSO in question (i.e. band-pass with high resonant-frequency: 0.2ms; band-pass with low resonant-frequency: 0.5ms). This increase in synaptic time course is performed to reflect the slower membrane time constant associated with neurons displaying little or no resonance at all. If, as expected, MSO neurons with slower intrinsic properties still segregate inputs from either ear on either dendrite then the increased electrotonic propagation of excitatory inputs somatopetally should lead to slower excitatory potentials at their somata. Where comparisons of the two MSO resonant classes are made, the same adapting AN inputs are presented to their respective CN stages, therefore reducing any potential confusion that might result from variability in the peripheral model's output.

Thresholding the linear-membrane model

Idealized thresholds for the linear-membrane model are implemented from Remme *et al.* (2014). A slope threshold (dv/dt) is used for all neuronal types. Threshold values for VCN cells are selected to produce good average firing rate without undermining the effects of synaptic depression. Threshold values for the MSO neurons are selected to obtain maximum dynamic range of ITD tuning functions for 200/600 Hz pure tone stimuli (where dynamic range is considered the difference between the maximum and minimum ITD-modulated spike rate). A refractory period of 1 ms is implemented as in Remme *et al.* (2014) for VCN neurons and MSO neurons with a high-resonant frequency whereas a longer refractory period of 2 ms is used for the MSO neurons with low-resonant frequency, reflecting longer refractory periods in neurons with lower densities of K_{LT} channels (Rothman & Manis, 2003b, 2003a, 2003c).

Calculating lateralisation using d' values

Lateralisation of a reverberant natural stimulus is judged in the linear model by calculating the mean spike rate difference between a population of 50 'left' and 50 'right' MSO neurons to a single stimulus presentation. d' values are calculated to quantify the degree of separation of 'left' and 'right' spike rates using the following equation:

$$d' = \frac{\mu_L - \mu_R}{\sqrt{\frac{1}{2}(\sigma_L^2 + \sigma_R^2)}}$$

Where $\mu_L$is the mean spike rate for the 'left' population of MSO neurons in 5-ms bins; $\mu_R$is the mean spike rate for the 'right' population of MSO neurons in 5-ms bins; σ_L is the standard deviation of the 'left' MSO population's spike rate and σ_R is the standard deviation of the 'right' population's spike rate (both calculated in 5-ms bins again).

The sign and magnitude of the d' allows judgement of whether the linear model lateralises the reverberant stimulus to the correct direction (negative d' values denote negative ITDs/left of midline; positive d' values denote positive ITDs/right of midline) and whether this lateralisation judgment is considered significant (d' > 1 or d' < -1 indicated a significant lateralisation in the right and left directions respectively).

ANOVA and T-tests tests for statistically significant differences in lateralisation

Two-way and three-way ANOVAs are implemented to compare the duration of correct lateralisations (time correctly lateralised = bin duration x number of bins where d' > 1) by the hemispheric model of SAM tones with non-zero ITDs in either the rising, peak or falling phases as well as the percentage of the phase including a non-zero ITD. A two-tailed T-test is also performed for 200-Hz carrier SAM tones to compare the effect of MSO neuronal type. For the speech stimulus, a paired two-tailed T-test is performed to compare the frequency of incorrect lateralisations (time incorrectly lateralised = bin duration x number of bins where d' <-1) in the linear model with either adapting or non-adapting AN inputs.

Author Contributions

AB and JMH contributed equally. Research concept: DM. Nonlinear modelling: AB. Linear modelling: JMH. Analysed and interpreted data: AB, JMH, DM, and MD. Drafted manuscript: AB and JMH. Revised manuscript: AB, DM, JMH, and MD.

BIBLIOGRAPHY

Adams, J. C., & Mugnaini, E. (1990). Immunocytochemical evidence for inhibitory and disinhibitory circuits in the superior olive. *Hearing Research*, *49*(1–3), 281–298. https://doi.org/10.1016/0378-5955(90)90109-3

Bondy, B., & Golding, N. L. (2018). Variations in intrinsic and synaptic properties in MSO neurons confer a spectrum of ITD sensitivity. *ARO Midwinter Meeting, San Diego, CA, USA*.

Bradley, J. S., Sato, H., & Picard, M. (2003). On the importance of early reflections for speech in rooms. *The Journal of the Acoustical Society of America*, *113*(6), 3233. https://doi.org/10.1121/1.1570439

Brown, A. D., Jones, H. G., Kan, A., Thakkar, T., Stecker, G. C., Goupell, M. J., & Litovsky, R. Y. (2015). Evidence for a neural source of the precedence effect in sound localization. *Journal of Neurophysiology*, *114*(5), 2991–3001. https://doi.org/10.1152/jn.00243.2015

Brughera, A., Dunai, L., & Hartmann, W. M. (2013). Human interaural time difference thresholds for sine tones: The high-frequency limit. *The Journal of the Acoustical Society of America*, *133*(5), 2839–2855. https://doi.org/10.1121/1.4795778

Burger, R. M., & Pollak, G. D. (2001). Reversible Inactivation of the Dorsal Nucleus of the Lateral Lemniscus Reveals Its Role in the Processing of Multiple Sound Sources in the Inferior Colliculus of Bats. *The Journal of Neuroscience*, *21*(13), 4830–4843. https://doi.org/10.1523/jneurosci.21-13-04830.2001

Cant, N. B., & Hyson, R. L. (1992). Projections from the lateral nucleus of the trapezoid body to the medial superior olivary nucleus in the gerbil. *Hearing Research*, *58*(1), 26–34. https://doi.org/10.1016/0378-5955(92)90005-8

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with 2 ears. *JASA*, (25), 975–979. https://doi.org/10.1121/1.1907229

Chung, Y., Delgutte, B., & Colburn, H. S. (2015). Modeling Binaural Responses in the Auditory Brainstem to Electric Stimulation of the Auditory Nerve. *JARO - Journal of the Association for Research in Otolaryngology*, *16*(1), 135–158. https://doi.org/10.1007/s10162-014-0492-6

Colburn, H. S., Chung, Y., Zhou, Y., & Brughera, A. (2009). Models of brainstem responses to bilateral electrical stimulation. *JARO - Journal of the Association for Research in Otolaryngology*, *10*(1), 91–110. https://doi.org/10.1007/s10162-008-0141-z

Couchman, K., Grothe, B., & Felmy, F. (2010). Medial Superior Olivary Neurons Receive Surprisingly Few Excitatory and Inhibitory Inputs with Balanced Strength and Short-Term Dynamics. *Journal of Neuroscience*, *30*(50), 17111–17121. https://doi.org/10.1523/JNEUROSCI.1760-10.2010

Devore, S., Ihlefeld, A., Hancock, K., Shinn-Cunningham, B., & Delgutte, B. (2009).

Accurate Sound Localization in Reverberant Environments Is Mediated by Robust Encoding of Spatial Cues in the Auditory Midbrain. *Neuron*, *62*(1), 123–134. https://doi.org/10.1016/j.neuron.2009.02.018

Dietz, M., Ewert, S. D., & Hohmann, V. (2009). Model-based direction estimation of concurrent speakers in the horizontal plane. *Journal of the Acoustical Society of America*, *125*(4), 2527.

Dietz, M., Marquardt, T., Salminen, N. H., & McAlpine, D. (2013). Emphasis of spatial cues in the temporal fine structure during the rising segments of amplitude-modulated sounds. *Proceedings of the National Academy of Sciences*, *110*(37), 15151–15156. Retrieved from www.pnas.org/cgi/doi/10.1073/pnas.1309712110

Dietz, M., Marquardt, T., Stange, A., Pecka, M., Grothe, B., & McAlpine, D. (2014). Emphasis of spatial cues in the temporal fine structure during the rising segments of amplitude-modulated sounds II: single-neuron recordings. *Journal of Neurophysiology*, *111*(10), 1973–1985. https://doi.org/10.1152/jn.00681.2013

Fitzpatrick, D. C., Kuwada, S., Batra, R., & Trahiotis, C. (1995). Neural responses to simple simulated echoes in the auditory brain stem of the unanesthetized rabbit. *J Neurophysiol*, *74*(6), 2469–2486. https://doi.org/10.1152/jn.1995.74.6.2469

Gai, Y., Doiron, B., Kotak, V., & Rinzel, J. (2009). Noise-Gated Encoding of Slow Inputs by Auditory Brain Stem Neurons With a Low-Threshold K+ Current. *Journal of Neurophysiology*, *102*(6), 3447–3460. https://doi.org/10.1152/jn.00538.2009

Gourévitch, B., & Brette, R. (2012). The impact of early reflections on binaural cues. *The Journal of the Acoustical Society of America*, *132*(1), 9–27. https://doi.org/10.1121/1.4726052

Haas, H. (1951). The influence of a single echo on the audibility of speech. *Acoustica*, *1*, 49–58.

Harper, N. S., & McAlpine, D. (2004). Optimal neural population coding of an auditory spatial cue. *Nature*, *430*(7000), 682–686. https://doi.org/10.1038/nature02768

Harper, N. S., Scott, B. H., Semple, M. N., & McAlpine, D. (2014). The neural code for auditory space depends on sound frequency and head size in an optimal manner. *PLoS ONE*, *9*(11). https://doi.org/10.1371/journal.pone.0108154

Hartung, K., & Trahiotis, C. (2001). Peripheral auditory processing and investigations of the “precedence effect” which utilize successive transient stimuli. *The Journal of the Acoustical Society of America*, *110*(3), 1505–1513. https://doi.org/10.1121/1.1390339

Hodgkin, A. L., & Huxley, A. F. (1952). A quantitative description of ion currents and its applications to conduction and excitation in nerve membranes. *J Physiol*, *117*, 500–544.

Hu, H., Ewert, S. D., McAlpine, D., & Dietz, M. (2017). Differences in the temporal course of interaural time difference sensitivity between acoustic and electric

hearing in amplitude modulated stimuli. *The Journal of the Acoustical Society of America*, *141*(3), 1862–1873. https://doi.org/10.1121/1.4977014

Ihlefeld, A., & Shinn-Cunningham, B. G. (2011). Effect of source spectrum on sound localization in an everyday reverberant room. *The Journal of the Acoustical Society of America*, *130*(1), 324–333. https://doi.org/10.1121/1.3596476

Johnson, D. H. (1980). The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. *The Journal of the Acoustical Society of America*, *68*, 1115–1122. https://doi.org/10.1121/1.384982

Joris, P. X., Carney, L. H., Smith, P. H., & Yin, T. C. (1994). Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency. *Journal of Neurophysiology*. https://doi.org/10.1152/jn.1994.71.3.1022

Joris, Philip X., & Yin, T. C. T. (1992). Responses to amplitude-modulated tones in the auditory nerve of the cat. *The Journal of the Acoustical Society of America*, *91*, 215–232. https://doi.org/10.1121/1.402757

Keine, C., Rubsamen, R., & Englitz, B. (2017). Signal integration at spherical bushy cells enhances representation of temporal structure but limits its range. *ELife*, *6*, 1–16. https://doi.org/10.7554/eLife.29639

Kidd, S. A., & Kelly, J. B. (1996). Contribution of the Dorsal Nucleus of the Lateral Lemniscus to Binaural Responses in the Inferior Colliculus of the Rat: Interaural Time Delays. *The Journal of Neuroscience*, *16*(22), 7390–7397. https://doi.org/10.1523/jneurosci.16-22-07390.1996

Klumpp, R. G., & Eady, H. R. (1956). Some Measurements of Interaural Time Difference Thresholds. *The Journal of the Acoustical Society of America*, *28*(5), 859–860. https://doi.org/10.1121/1.1908493

Liebenthal, E., & Pratt, H. (2002). Human auditory cortex electrophysiological correlates of the precedence effect: Binaural echo lateralization suppression. *The Journal of the Acoustical Society of America*, *106*(1), 291–303. https://doi.org/10.1121/1.427057

Litovsky, R. Y., Colburn, H. S., Yost, W. A., & Guzman, S. J. (1999). The precedence effect. *J. Acoust. Soc. Am.*, *106*(4), 1633–1654.

Litovsky, R.Y., & Delgutte, B. (2002). Neural Correlates of the Precedence Effect in the Inferior Colliculus: Effect of Localization Cues. *Journal of Neurophysiology*, *87*(2), 976–994. https://doi.org/10.1016/j.pediatrneurol.2013.11.021

Litovsky, R Y, & Yin, T. C. (1998). Physiological studies of the precedence effect in the inferior colliculus of the cat. I. Correlates of psychophysics. *Journal of Neurophysiology*, *80*(3), 1285–1301. https://doi.org/10.1121/1.423072

Litovsky, Ruth Y., & Yin, T. C. T. (1998). Physiological Studies of the Precedence Effect in the Inferior Colliculus of the Cat. II. Neural Mechanisms. *Journal of*

Neurophysiology, *80*(3), 1302–1316. https://doi.org/10.1152/jn.1998.80.3.1302

Manis, P. B., & Marx, S. O. (1991). Outward currents in isolated ventral cochlear nucleus neurons. *The Journal of Neuroscience*, *11*(9), 2865–2880. https://doi.org/10.1523/JNEUROSCI.2541-12.2013

Mathews, P. J., Jercog, P. E., Rinzel, J., Scott, L. L., & Golding, N. L. (2010). Control of submillisecond synaptic timing in binaural coincidence detectors by K v 1 channels. *Nature Neuroscience*, *13*(5), 601–609. https://doi.org/10.1038/nn.2530

McAlpine, D., Jiang, D., & Palmer, A. R. (2001). A neural code for low-frequency sound localization in mammals. *Nature Neuroscience*, *4*(4), 396–401. https://doi.org/10.1038/86049

Mikiel-Hunter, J., Kotak, V., & Rinzel, J. (2016). High-Frequency Resonance in the Gerbil Medial Superior Olive. *PLoS Computational Biology*, *12*(11), 1–23. https://doi.org/10.1371/journal.pcbi.1005166

Oleskevich, S., Clements, J., & Walmsley, B. (2000). Release probability modulates short-term plasticity at a rat giant terminal. *Journal of Physiology*, *524*(2), 513–523. https://doi.org/10.1111/j.1469-7793.2000.00513.x

Pecka, M., Zahn, T. P., Saunier-Rebori, B., Siveke, I., Felmy, F., Wiegrebe, L., … Grothe, B. (2007). Inhibiting the Inhibition: A Neuronal Network for Sound Localization in Reverberant Environments. *Journal of Neuroscience*, *27*(7), 1782–1790. https://doi.org/10.1523/JNEUROSCI.5335-06.2007

Remme, M. W. H., Donato, R., Mikiel-Hunter, J., Ballestero, J. A., Foster, S., Rinzel, J., & McAlpine, D. (2014). Subthreshold resonance properties contribute to the efficient coding of auditory spatial cues. *Proceedings of the National Academy of Sciences*, *111*(22), E2339–E2348. https://doi.org/10.1073/pnas.1316216111

Rhode, W. S. (1976). A test for the significance of the mean direction and the concentration parameter of a circular distribution. *Madison: University of Wisconsin Dept. of Neurophysiology Report.* Retrieved from http://www.neurophys.wisc.edu/comp/docs/not011/not011.html

Rhode, W. S., & Greenberg, S. (1994). Encoding of amplitude modulation in the cochlear nucleus of the cat. *Journal of Neurophysiology*, *71*(5), 1797–1825. https://doi.org/10.1152/jn.1994.71.5.1797

Rothman, J. S., & Manis, P. B. (2003a). Differential Expression of Three Distinct Potassium Currents in the Ventral Cochlear Nucleus. *Journal of Neurophysiology*, *89*(6), 3070–3082. https://doi.org/10.1152/jn.00125.2002

Rothman, J. S., & Manis, P. B. (2003b). Kinetic Analyses of Three Distinct Potassium Conductances in Ventral Cochlear Nucleus Neurons. *Journal of Neurophysiology*, *89*(6), 3083–3096. https://doi.org/10.1152/jn.00126.2002

Rothman, J. S., & Manis, P. B. (2003c). The Roles Potassium Currents Play in Regulating the Electrical Activity of Ventral Cochlear Nucleus Neurons. *Journal of Neurophysiology*, *89*(6), 3097–3113. https://doi.org/10.1152/jn.00127.2002

Rudnicki, M., & Hemmert, W. (2017). High Entrainment Constrains Synaptic Depression Levels of an In vivo Globular Bushy Cell Model. *Frontiers in Computational Neuroscience*, *11*(March), 1–11. https://doi.org/10.3389/fncom.2017.00016

Scott, L. L., Mathews, P. J., & Golding, N. L. (2010). Perisomatic Voltage-Gated Sodium Channels Actively Maintain Linear Synaptic Integration in Principal Neurons of the Medial Superior Olive. *Journal of Neuroscience*, *30*(6), 2039–2050. https://doi.org/10.1523/JNEUROSCI.2385-09.2010

Scott, Luisa L., Hage, T. A., & Golding, N. L. (2007). Weak action potential backpropagation is associated with high-frequency axonal firing capability in principal neurons of the gerbil medial superior olive. *Journal of Physiology*, *583*(2), 647–661. https://doi.org/10.1113/jphysiol.2007.136366

Shinn-Cunningham, B. G., Zurek, P. M., Durlach, N. I., & Clifton, R. K. (1995). Cross-frequency interactions in the precedence effect. *The Journal of the Acoustical Society of America*, *98*(1), 164–171. https://doi.org/10.1121/1.413752

Smith, P. H., Joris, P. X., Carney, L. H., & Yin, T. C. T. (1991). Projections of physiologically characterized globular bushy cell axons from the cochlear nucleus of the cat. *Journal of Comparative Neurology*, *304*(3), 387–407. https://doi.org/10.1002/cne.903040305

Stimberg, M., Goodman, D., Benichoux, V., & Brette, R. (2013). Brian 2-the second coming: spiking neural network simulation in Python with code generation. *Bmc …*, *14*(Suppl 1), P38. https://doi.org/10.1186/1471-2202-14-S1-P38

Svirskis, G., Kotak, V., Sanes, D., & Rinzel, J. (2002). Enhancement of signal-to-noise ratio and phase locking for small inputs by a low-threshold outward current in auditory neurons. *Journal of Neuroscience*, *22*(24), 11019–11025.

Svirskis, G., Kotak, V., Sanes, D., & Rinzel, J. (2004). Sodium Along With Low-Threshold Potassium Currents Enhance Coincidence Detection of Subthreshold Noisy Signals in MSO Neurons. *Journal of Neurophysiology*, *91*(6), 2465–2473. https://doi.org/10.1152/jn.00717.2003

Tollin, D. J., Populin, L. C., & Yin, T. C. T. (2004). Neural Correlates of the Precedence Effect in the Inferior Colliculus of Behaving Cats. *Journal of Neurophysiology*, *92*(6), 3286–3297. https://doi.org/10.1152/jn.00606.2004

Traer, J., & McDermott, J. H. (2016). Statistics of natural reverberation enable perceptual separation of sound and space. *Proceedings of the National Academy of Sciences*, *113*(48), E7856–E7865. https://doi.org/10.1073/pnas.1612524113

Wallach, H., Newman, E. B., & Rosenzweig, M. R. (1949). The precedence effect in sound localization. *Am J Psychol*, *62*(3), 315–336.

Wang, Y., & Manis, P. B. (2008). Short-Term Synaptic Depression and Recovery at the Mature Mammalian Endbulb of Held Synapse in Mice. *Journal of Neurophysiology*, *100*(3), 1255–1264. https://doi.org/10.1152/jn.90715.2008

Xia, J., Brughera, A., Colburn, H. S., & Shinn-Cunningham, B. (2010). Physiological and psychophysical modeling of the precedence effect. *JARO - Journal of the Association for Research in Otolaryngology*, *11*(3), 495–513. https://doi.org/10.1007/s10162-010-0212-9

Yang, H., & Xu-Friedman, M. A. (2009). Impact of Synaptic Depression on Spike Timing at the Endbulb of Held. *Journal of Neurophysiology*, *102*(3), 1699–1710. https://doi.org/10.1152/jn.00072.2009

Yin, C. T., & Chan, J. C. M. (1990). Interaural time sensitivity in medial superior olive of cat. *Journal of Neurophysiology*, *64*(2), 465–488. https://doi.org/10.1152/jn.1990.64.2.465

Yin, T. C. (1994). Physiological correlates of the precedence effect and summing localization in the inferior colliculus of the cat. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *14*(9), 5170–5186. https://doi.org/10.1523/JNEUROSCI.14-09-05170.1994

Zhou, Y., Carney, L. H., & Colburn, H. S. (2005). A Model for Interaural Time Difference Sensitivity in the Medial Superior Olive: Interaction of Excitatory and Inhibitory Synaptic Inputs, Channel Dynamics, and Cellular Morphology. *Journal of Neuroscience*, *25*(12), 3046–3058. https://doi.org/10.1016/j.physa.2016.12.040

Zilany, M. S. A., Bruce, I. C., & Carney, L. H. (2014). Updated parameters and expanded simulation options for a model of the auditory periphery. *The Journal of the Acoustical Society of America*, *135*(1), 283–286. https://doi.org/10.1121/1.4837815

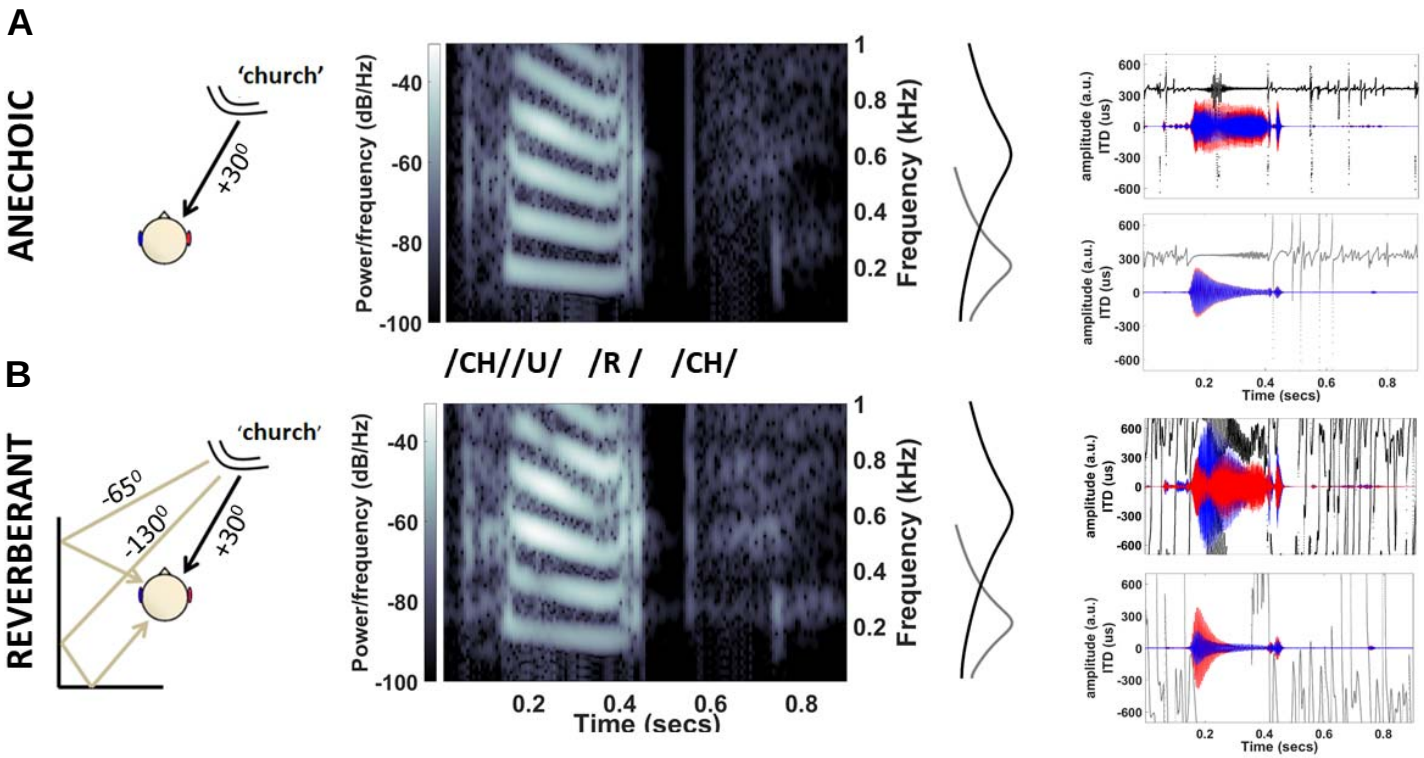


Figure 1. Reverberant copies of direct speech produce confounding binaural cues. (A) Anechoic speech stimulus 'church', spoken by female talker at 70 dB located +30° to the right and front of listener. Spectrogram of 'dry' speech at left ear plots signal energy <1 kHz. Gammatone filters centred at 200 Hz (light grey) and 600 Hz (dark grey) used to calculate waveforms at left (blue) and right (red) ears and instantaneous ITDs (light/dark grey) for both frequencies (restricted to ±700us human physiological range). ITDs for dry speech stimulus were consistently close to +363 µs. (**B) Reverberant speech stimulus** including direct speech plus two simulated reflections from listener's left, the first from -65° delayed by 4 ms, and the second from -130° delayed by 8 ms. Reverberant copies generated increased energy at left ear (the generally brighter spectrogram, esp. in the 600-Hz channel). Additional reverberant energy also extends into the quiet gap between vowel and the second consonant (0.45-0.55 ms), generating rapidly varying instantaneous ITDs (also restricted to ±700us), along with conflicting ILD cues (the latter esp. in the 600-Hz channel).

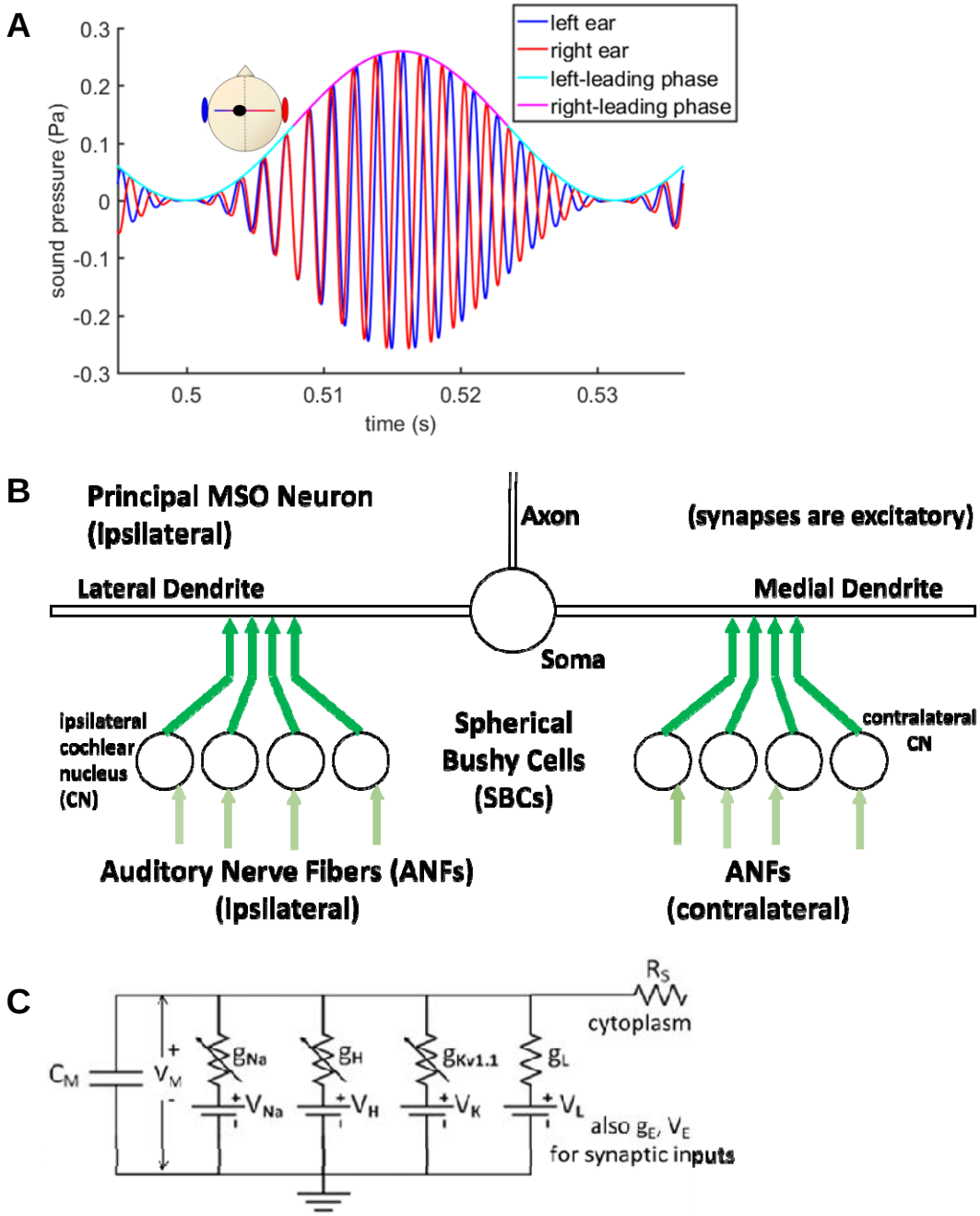
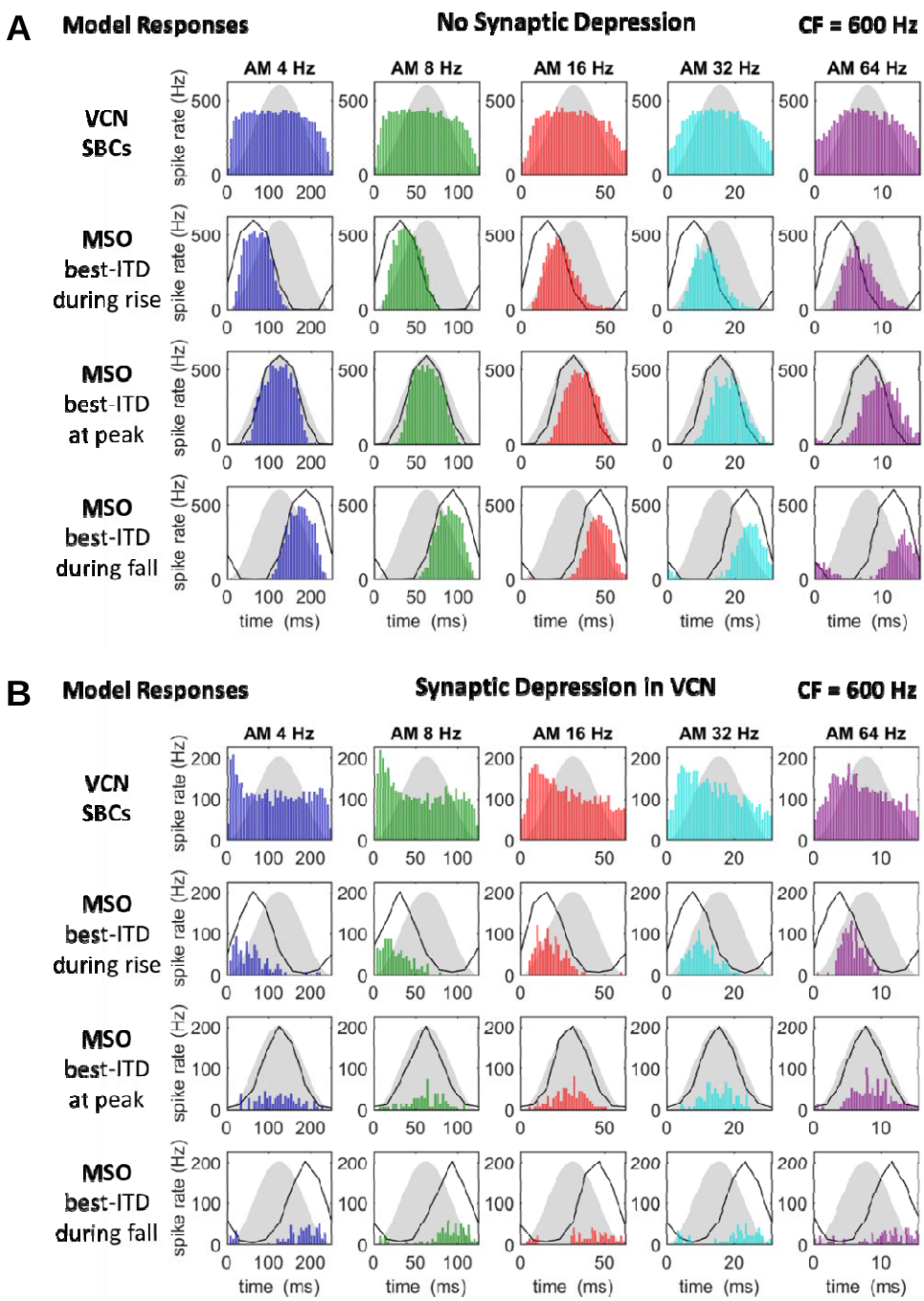


Figure 2. (A) Amplitude-modulated binaural beats (AMBBs) (Dietz et al., 2013) in which AM rate is equal to binaural-beat frequency (difference between frequencies at right and left ears). Presented via headphones, interaural phase difference (IPD) cycles through 360° at the same rate as AM. *Start-IPD* (the IPD at zero amplitude) is a free parameter. Shown with right-carrier 616 Hz, left-carrier 584 Hz, AM 32 Hz, and *start-IPD* 270° (right channel trailing by 90°), resulting in zero IPD at the midpoint of rising-amplitude. **(B) Nonlinear brainstem model** for MSO neuron and its excitatory inputs with synaptic depression in the CN. AMBBs drive a model for bilateral AN-fibres (Zilany et al., 2014) projecting individually depressing synapses (Rudnicki and Hemmert, 2017; u = 0.5) onto model SBCs of the CN (Rothman and Manis, 2003), which in turn project bilaterally to the model MSO neuron (best-ITD = 0). **(C) Sub-compartment** in the nonlinear brainstem model (see Methods).

Figure 3. AMBB cycle histograms for nonlinear model brainstem neurons acoustically stimulated by sound frequencies centred at CF **600 Hz (A) Without synaptic depression**, AN to SBC synapses were strongly supra-threshold. Model SBCs adapted slightly. Model MSO neuron responded to best ITD strongly at peak sound energy, and systematically less strongly during rising and falling energy (3-way comparisons, $p < 10^{-9}$). The synchrony index—a measure of temporal alignment of spikes over the AM cycle—ranged from 0.133 at 4 and 64 Hz to 0.161 at 16 Hz for model SBCs, and from 0.595 (64 Hz, best-ITD at peak) to 0.794 (16 Hz, best-ITD during fall) for the model MSO neuron (Rayleigh statistic, $p < 0.001$ in all cases). **(B) With synaptic depression in the VCN**, AN to SBC synapses were supra-threshold at full strength. Model SBCs adapted. Model MSO neuron responded to best ITD strongly during rising energy; weakly to moderately at peak; and weakly during falling, energy (3-way comparisons, $p \leq 0.003$). Synchrony-indices to AM for model SBCs ranged from 0.0485 at 4 Hz to 0.174 at 32 Hz; model MSO neuron from 0.450 (64 Hz, best-ITD during fall) to 0.814 (64 Hz, best-ITD during rise) (Rayleigh statistic, $p < 0.01$ for model SBCs at 4-Hz AM, otherwise $p < 0.001$).

Grey silhouettes: AM envelope. Black lines: static IPD functions. Best-ITD = 0. Best-ITD during rise: *start-IPD* = 270°. Best-ITD at peak: *start-IPD* = 180°. Best-ITD during fall: *start-IPD* = 90°.



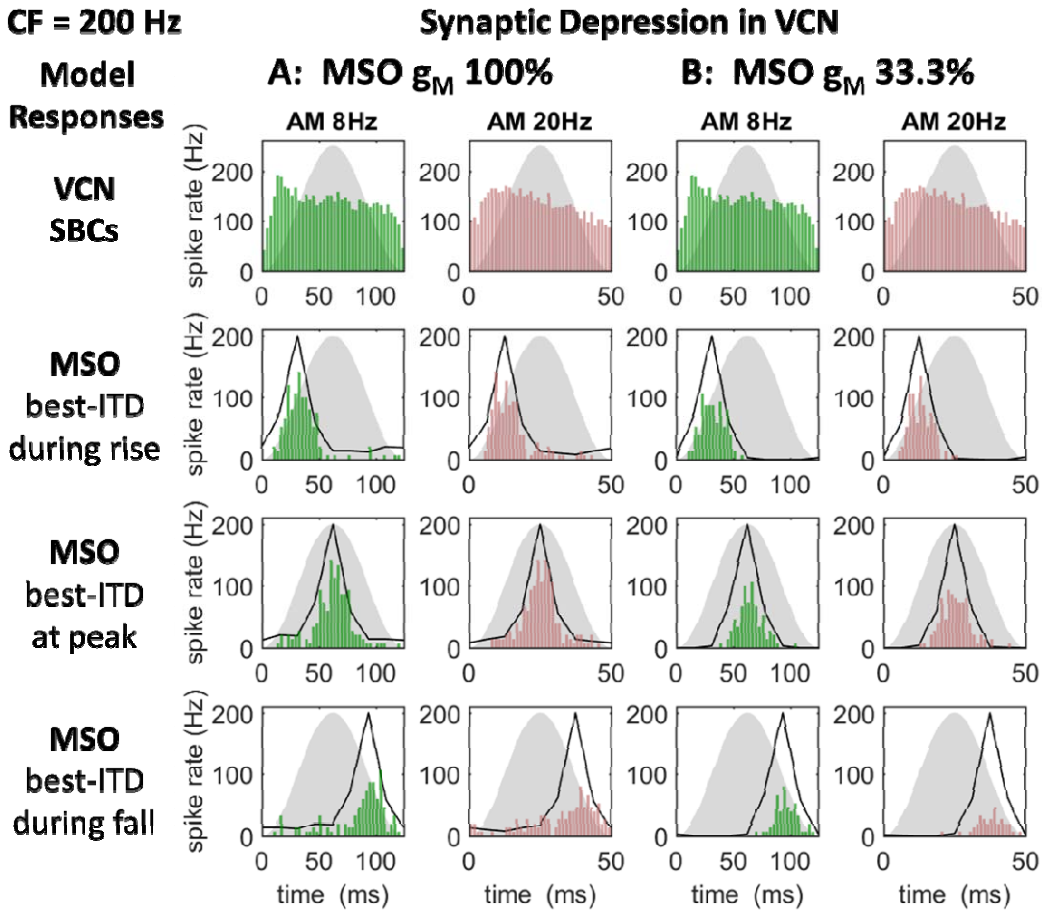


Figure 4. AMBB cycle histograms for nonlinear model brainstem neurons stimulated by frequencies centred at CF **200 Hz**. Same synaptic depression and model SBCs as in Fig. 3B, adapted less at 200 Hz. **(A) Same fast model MSO neuron** as in Fig. 3B: spike counts were not significantly different for best-ITD during rising vs. peak energy, and were higher for best-ITD at peak energy vs. falling energy ($p < 0.01$). Synchrony-indices to AM for model SBCs ranged from 0.0762 at 8 Hz to 0.109 at 20 Hz; model MSO neuron from 0.433 (20 Hz, best-ITD during fall) to 0.807 (8 Hz, best-ITD during rise) (Rayleigh statistic, $p < 0.001$). **(B) Moderate-speed model MSO neuron:** spike counts were not significantly different for best-ITD during rising vs. peak energy, and were lower for best ITD during falling energy ($p < 0.005$). Grey silhouettes: AM envelope. Black lines: static IPD functions. Best ITD = 0. Best-ITD during rise: *start-IPD* = 270°. Best ITD at peak: *start-IPD* = 180°. Best ITD during fall *start-IPD* = 90°. Synchrony-indices to AM for the model MSO neuron ranged from 0.825 (20 Hz, best-ITD during fall) to 0.888 (20 Hz, best-ITD during rise) (Rayleigh statistic, $p < 0.001$)

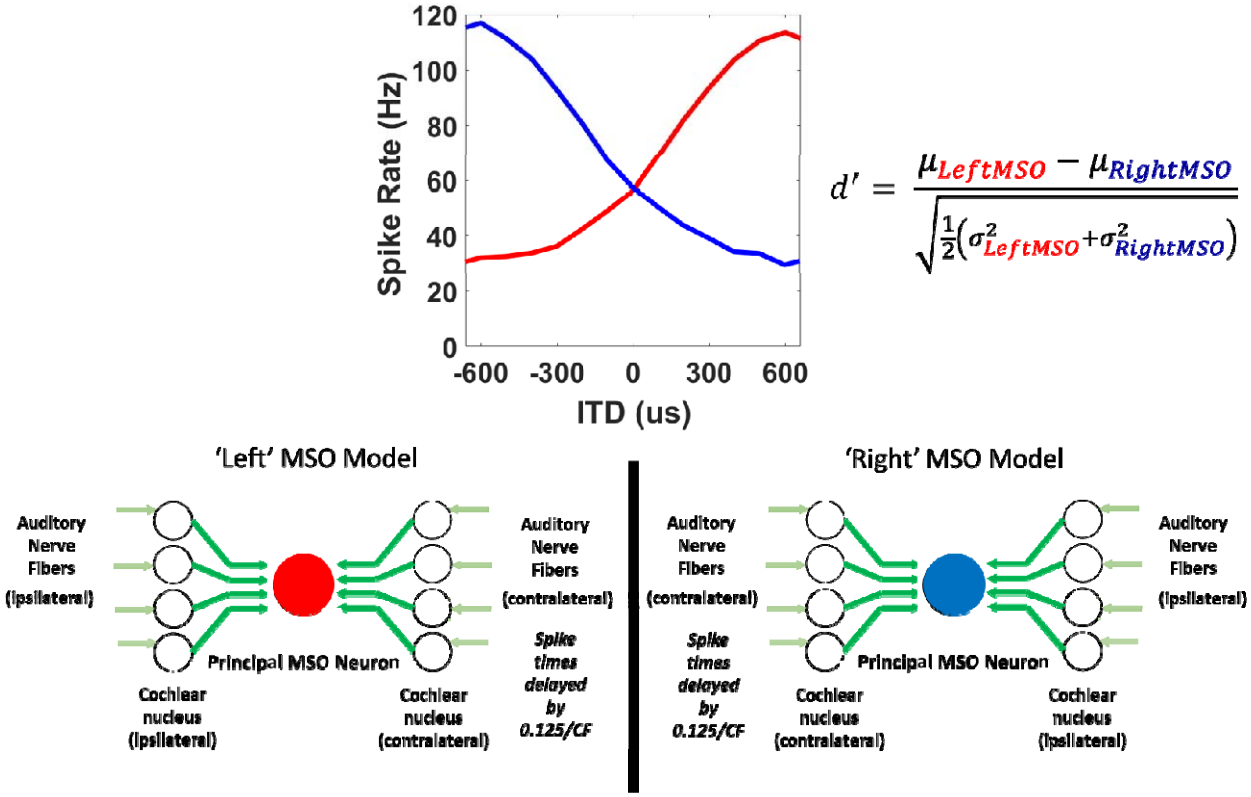


Figure 5. Overview of the hemispheric-difference model consisting of two MSO populations (red—left MSO and blue—right MSO), each containing fifty linear, single-compartment model neurons, with either fast or slow resonant properties (Remme *et al*., 2014). Each MSO neuron model receives four excitatory inputs 'bilaterally' from linear, single-compartment SBC model neurons in the CN. Four independently-simulated, individually-depressing and medium-spontaneous-rate AN-fibres (Zilany *et al*., 2014) provide excitatory drive to each SBC model neuron.

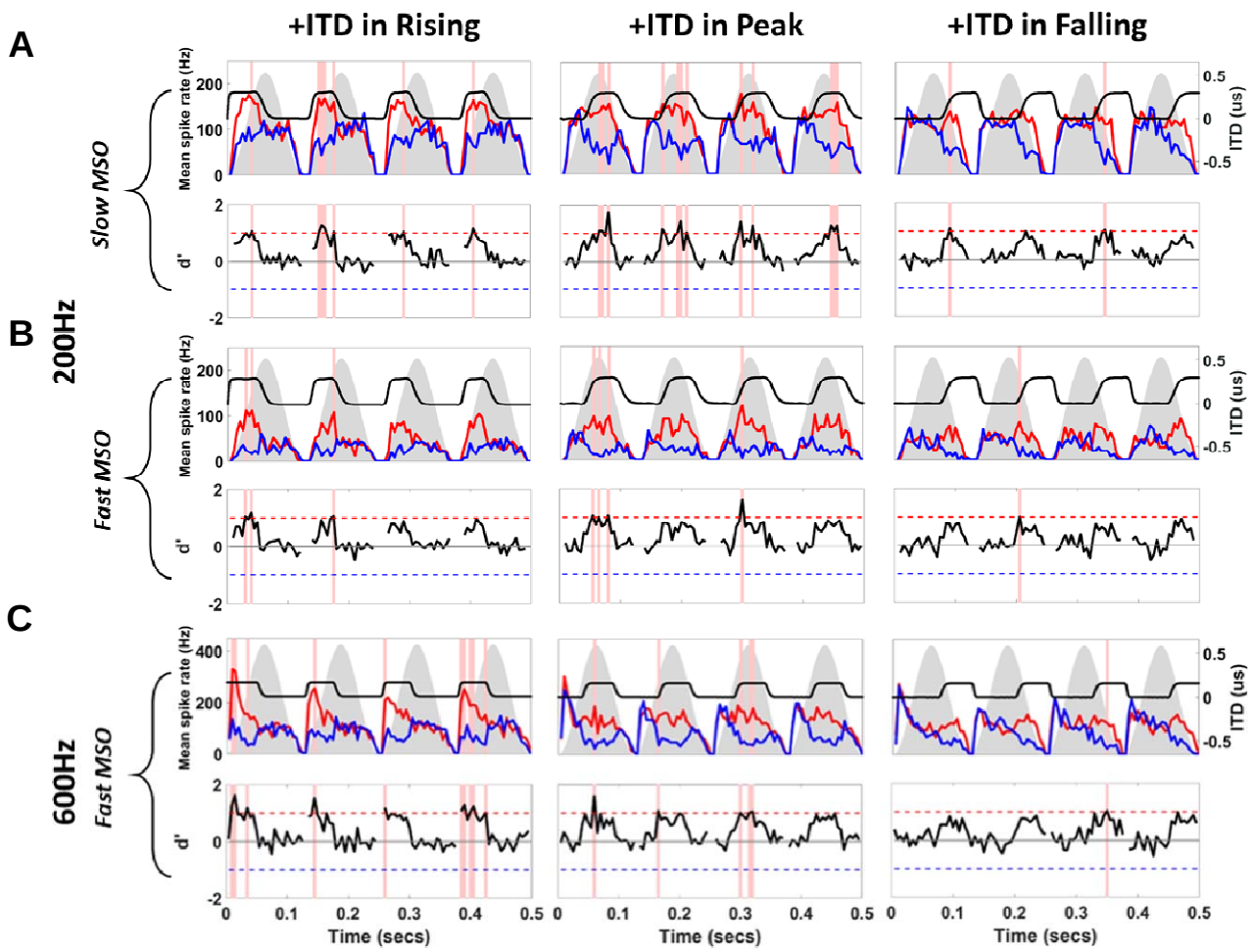


Figure 6. AM lateralisations by linear, hemispheric-difference model at 200Hz and 600Hz. (A&B) At 200 Hz, both Slow (A) and Fast (B) MSO models correctly lateralised +300 µs ITDs (top row, black) to the right (d'>1 in any 5ms bin (red vertical bands) where d' (bottom rows, black) calculated from mean spike rates of left (top rows, red) and right (top rows, blue) MSO populations) most often at peak (middle column) of 8-Hz AM cycle (top row, grey silhouettes) compared to rising (left column) and falling (right column) phases (Slow MSO: Rising, 7 bins; Peak, 12 bins; Falling, 2 bins. Fast MSO: Rising, 3 bins; Peak, 4 bins; Falling, 1 bin). Same adapting AN inputs used for both MSO neuronal types, therefore more correct lateralisations overall by Slow MSO is a result of its slower integration. **At 600 Hz (C)**, introducing +150 us ITD (top row, black) to 8 Hz AM (top row, grey silhouettes) produced strong onset ITD sensitivity in Fast MSO towards rising phase (left column) that decreased across AM cycle (Fast MSO: Rising, 12 bins; Peak, 5 bins; Falling, 1 bin).

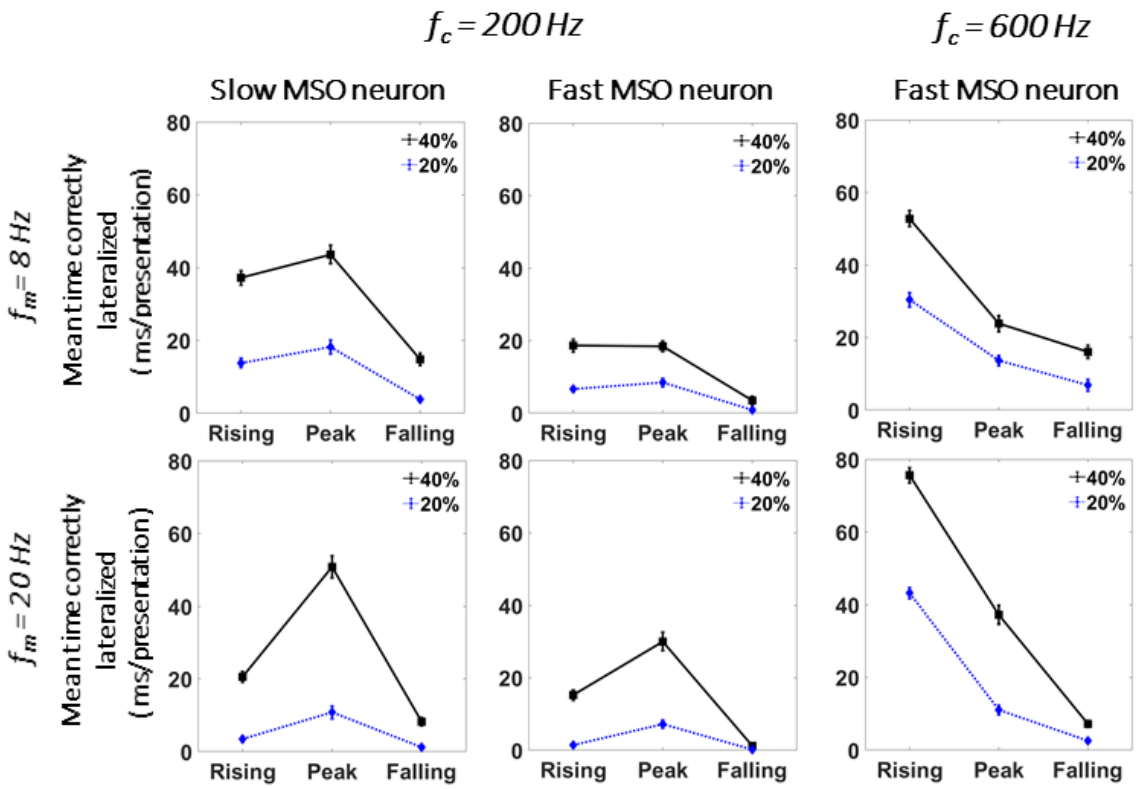
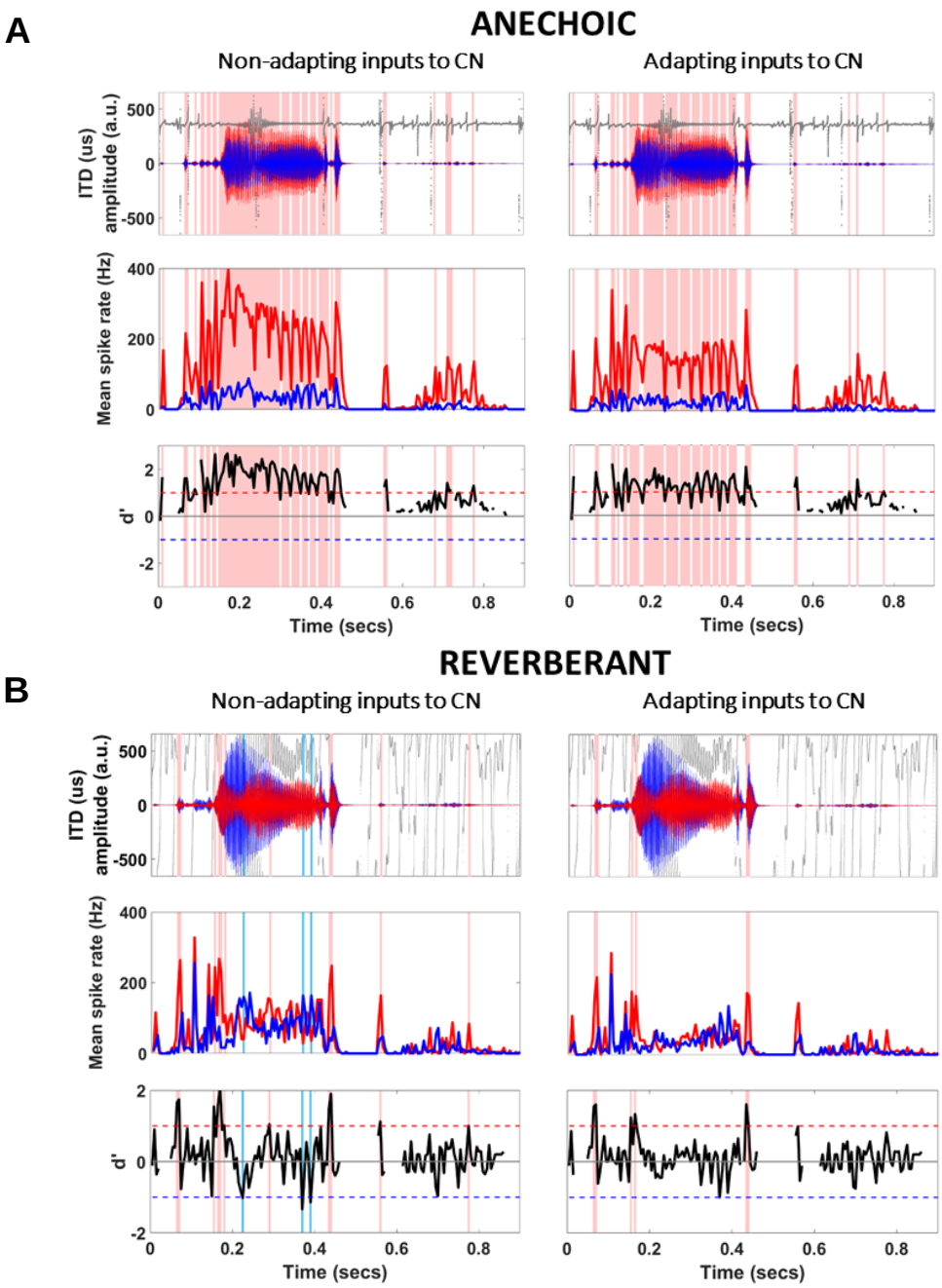


Figure 7. Summarizing correct lateralisations of SAM stimuli by the hemispheric-difference model. Mean (±SEM) time correctly lateralised (number of correct lateralisations * bin duration) per presentation for 25 presentations to either Slow MSO neurons at 200 Hz (left column) or Fast MSO neurons at 200 Hz and 600 Hz (middle and right columns respectively). Modulation rates of 8 Hz (top row) and 20 Hz (bottom row) were tested. ITDs were inserted into either 20% (blue dotted) or 40% (black solid) of the AM cycle during Rising, Peak, and Falling energy. At 600 Hz, the Fast MSO displays onset dominance of ITD weighting that decreased significantly across AM cycle (Two-way ANOVA, interaction of AM cycle phase x Percentage ITD insertion: 8-Hz AM, $F(2,48) = 8.310$, $p = 0.001$; 20-Hz AM, $F(2,48) = 50.098$, $p = 0.0001$). At 200 Hz, the Fast MSO showed equal or better ITD sensitivity at Peak compared to Rising phase (Two-way ANOVA, interaction of AM cycle phase x Percentage ITD insertion: 8-Hz AM, $F(2,48) = 7.83$, $p = 0.001$; 20-Hz AM, $F(2,48) = 36.01$, $p = 0.0001$). Slow MSO generated more correct lateralisations than Fast MSO (Two-way ANOVA, interaction of AM cycle phase x Neuron Type: 8-Hz AM, $F(2,48) = 18.82$, $p = 0.0001$; 20-Hz AM, $F(2,48) = 13.12$, $p = 0.0001$) with equal or augmented ITD sensitivity at Peak for 40% non-zero ITD (Paired two-tailed T-test, Mean Peak – Rising, N = 25: 8-Hz AM, $t(24) = 2.21$, $p = 0.037$; 20-Hz AM, $t(24) = 2.21$, $p = 2.37 \times 10^{-4}$). Correct lateralisations also increased as the percentage of AM phase containing non-zero ITD was raised from 20% to 40% (Three-way ANOVA, main effect of Percentage ITD insertion: Fast MSO, $F(1,24) = 418.78$, $p = 0.0001$; Slow MSO, $F(1,24) = 285.622$, $p = 0.0001$)



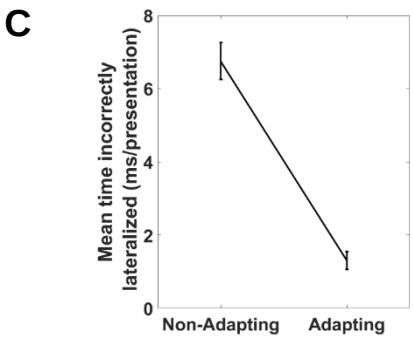


Figure 8. Lateralisation of 'anechoic' and 'reverberant' speech stimuli by the hemispheric-difference model (fast MSO, 600-Hz channel) with and without adapting auditory nerve inputs. (A) Anechoic stimulus is correctly lateralised to the right (pink vertical bands) for long periods independently of non-adapting (left column) or adapting (right column) AN inputs. Dry speech stimulus gammatone-filtered at 600 Hz with instantaneous ITDs (top row), mean firing rates of left (middle row, red) and right (middle row, blue) MSOs and d' neuro-metric (bottom row) are displayed. **(B) Reverberant stimulus** requires adapting AN inputs (right column) for correct lateralisations alone (pink vertical bands). Non-adapting AN inputs to CN produced both correct (left column, pink vertical bands) and incorrect lateralisations (left column, light blue vertical bands) due to reverberant stimulus' highly variable ITDs and confounding ILDs (top row). The inclusion of synaptic depression from AN to CN stage (same AN simulations used for both adapting/non-adapting conditions) removes all incorrect lateralisations. Remaining correct lateralisations correspond with stimulus onsets (right column, pink vertical bands). **(C) Quantifying incorrect lateralisations with non-adapting and adapting AN inputs** (same AN simulations for both) over a hundred presentations of the reverberant speech stimulus. Synaptic depression produces a fivefold decrease in incorrect lateralisations (N = 100, t(99) = 9.51, $p = 1.31 \times 10^{-15}$, two-tailed paired T-test).

Table 1. Model MSO neuron parameters with membrane conductance at full value

Parameter	Dendrites (2)	Soma (spherical)	Axon
Temperature (°C)	37	37	37
Number of sub-compartments	20	2	51
Diameter (μm)	3.5	30	2
Length (μm)	150	30	400
Resistivity, R_S (ohm x cm)	150	0 (iso-potential)	150
C_M (μF/cm^2)	1	1	1
V_{KLT} (mV)	-106	-106	-106
V_{Na} (mV)	n/a	-62.1	-62.1
V_H (mV)	-43	-43	-43
V_L (mV)	-60	-60	-65
$gmax_{KLT}$ (S/cm^2)	0.00132	0.0324	0.0595
$gmax_{Na}$ (S/cm^2)	0 (none)	0.0432	0.25
$gmax_H$ (S/cm^2)	0.00066	0.0216	0.0025
g_L (S/cm^2)	0.00005	0.00005	0.00005
V_E (mV)	0	n/a	n/a
$V_{AP\text{-}THRESHOLD}$ (mV, set)	n/a	n/a	-20
V_{REST} (mV)	-60.53	-60.52	-64.35

Table 2. Each column gives parameter values for one of the four cases explored. Adjusted parameters were CF, VCN synaptic depression, MSO membrane conductance (in soma and dendrites), and MSO synaptic strength:

Results shown in	Fig. 3A	Fig. 3B	Fig. 4A	Fig. 4B
CF (Hz)	600	600	200	200
Model SBC excitatory synaptic parameters:				
Full Synaptic Strength, Δg_E (nS)	400	83	83	83
Synaptic Time Constant, τ_E (ms)	0.2	0.2	0.2	0.2
Depression, u	0 (none)	0.5	0.5	0.5
Recovery Time Constant (ms)	n/a	25	25	25
Model MSO membrane conductance factor and excitatory synaptic parameters:				
$g_{MEMBRANE}$ Factor	1	1	1	1/3
Membrane Time Constant (ms)	0.33	0.33	0.33	1.0
Synaptic Strength, Δg_E (nS)	21	36	36	18
Synaptic Time Constant, τ_E (ms)	0.4	0.4	0.4	0.4