Temporal plasticity in auditory cortex improves neural discrimination of speech sounds

Crystal T. Engineer^{* a,b}, Jai A. Shetake^{* a}, Navzer D. Engineer^{a,c}, Will A. Vrana^a, Jordan T. Wolf^a, Michael P. Kilgard^{a,b}

^a School of Behavioral and Brain Sciences, The University of Texas at Dallas, 800 West Campbell Road BSB11, Richardson, TX 75080, United States

^b Texas Biomedical Device Center, The University of Texas at Dallas, 800 West Campbell Road BSB11, Richardson, TX 75080, United States

^c MicroTransponder Inc., 2802 Flintrock Trace Suite 225, Austin, TX 78738, United States

* These authors contributed equally.

Corresponding author:

Crystal T. Engineer

novitski@utdallas.edu

The University of Texas at Dallas

800 West Campbell Road, BSB11

Richardson, TX 75080

Phone: 972-883-7246

20 **Abstract**

21 *Background*

22 Many individuals with language learning impairments exhibit temporal processing deficits and
23 degraded neural responses to speech sounds. Auditory training can improve both the neural and
24 behavioral deficits, though significant deficits remain. Recent evidence suggests that vagus nerve
25 stimulation (VNS) paired with rehabilitative therapies enhances both cortical plasticity and
26 recovery of normal function.

27 *Objective/Hypothesis*

28 We predicted that pairing VNS with rapid tone trains would enhance the primary auditory cortex
29 (A1) response to unpaired novel speech sounds.

30 *Methods*

31 VNS was paired with tone trains 300 times per day for 20 days in adult rats. Responses to
32 isolated speech sounds, compressed speech sounds, word sequences, and compressed word
33 sequences were recorded in A1 following the completion of VNS-tone train pairing.

34 *Results*

35 Pairing VNS with rapid tone trains resulted in stronger, faster, and more discriminable A1
36 responses to speech sounds presented at conversational rates.

37 *Conclusion*

38 This study extends previous findings by documenting that VNS paired with rapid tone trains
39 altered the neural response to novel unpaired speech sounds. Future studies are necessary to
40 determine whether pairing VNS with appropriate auditory stimuli could potentially be used to

41 improve both neural responses to speech sounds and speech perception in individuals with
42 receptive language disorders.

43

44 **Keywords** Speech therapy; aphasia; autism; dyslexia; vagal nerve stimulation

45

46

Introduction

The ability to follow the rapid spectrotemporal transitions in speech is necessary for effective speech processing. Speech comprehension ability can be predicted by the temporal processing capability of the auditory cortex. For example, a decrease in the ability of auditory cortex to reliably respond to every presented sound in a rapid train of sounds is associated with poor comprehension of rapidly presented speech [1,2]. Deficits in the temporal processing of rapid sounds contribute to speech processing problems in both developmental disorders (such as dyslexia and autism) and acquired disorders (such as aphasia) [3–6]. Improvements in temporal processing are associated with improved speech processing. Auditory perceptual training can improve both the temporal following capacity of auditory cortex neurons and the accuracy of behavioral judgments [3,7–13]. Given that temporal deficits often persist even after extensive perceptual training [14,15], a method to further enhance training-induced neural plasticity could be beneficial for therapy in a variety of speech processing disorders.

Precisely timed neuromodulator release paired with sound presentation enables large-scale plasticity in the auditory cortex representation of the paired sound. Pairing nucleus basalis stimulation or vagus nerve stimulation (VNS) with a tone increases the auditory cortex response to the paired tone [16–18]. Similarly, pairing VNS with 5 pps (pulses per second) or 15 pps tone trains decreases or increases the auditory cortex response to rapidly presented sounds [19]. Pairing VNS with speech sounds strengthens the auditory cortex response to the paired speech sounds [20]. These findings indicate that VNS-sound pairing therapy may represent a method to manipulate temporal processing and improve speech processing.

In this study, we investigated how alterations in the temporal response properties of auditory cortex neurons affect the neural representation of rapid speech. We paired VNS with 5

pps or 15 pps tone trains and documented the primary auditory cortex (A1) response to speech sounds after 20 days of VNS pairing. We hypothesized that VNS paired with 15 pps tone trains would increase the response strength and decrease the response latency to speech sounds.

Materials and methods

Nineteen female Sprague Dawley rats were used in these experiments. Thirteen rats received vagus nerve stimulation paired with tone trains with a repetition rate of either 15 pulses per second, pps ($n = 7$ rats) or 5 pps ($n = 6$ rats). The remaining 6 rats were experimentally naïve controls that did not experience VNS or sound presentation. Our previous studies revealed large behavioral and neurophysiological differences when VNS was paired with sensory or motor events. However, there were no differences between experimentally naïve rats and 1) rats that received VNS which was not paired with an event or 2) rats that were exposed to sensory or motor events without VNS pairing [17,19,21–23]. As a result, the control rats in the current study did not experience exposure to sounds or exposure to VNS. The University of Texas at Dallas Institutional Animal Care and Use Committee approved all protocols and surgical procedures.

Vagus nerve surgery

Each rat was implanted with a custom made platinum iridium bipolar cuff electrode around the left cervical vagus nerve. All surgical procedures were identical to the procedures in previous studies [17–19]. Rats were anesthetized with sodium pentobarbital anesthesia (50 mg/kg), and received supplemental doses of dilute pentobarbital as needed to maintain areflexia. A bipolar cuff electrode was wrapped around the left vagus nerve and tunneled subcutaneously

to a headcap connector located on top of the skull. Rats had a week of recovery from surgery, and were given oral amoxicillin and carprofen.

Sound stimuli

The tone trains paired with VNS in this study were identical to the trains used in previous studies [19,24]. The tone repetition rate was either 5 pps or 15 pps. Each train was made up of six tones, and each tone was 25 ms in duration. The carrier frequency of the tones that made up each train was one of seven frequencies distributed across the rat hearing range (1.3, 2.2, 3.7, 6.3, 10.6, 17.8, or 29.9 kHz). The carrier frequency of the tones in each train was randomized from trial to trial but the repetition rate was always the same for an individual rat.

The speech sounds presented during the terminal neurophysiology component of the study were words presented in a CVC (consonant-vowel-consonant) context and were identical to the sounds used in previous studies [25–28]. The words ‘bad’, ‘dad’, ‘sad’, ‘wad’, and ‘yad’ were spoken by a female native English speaker, and will be referred to as “speech sounds”. To better match the rat hearing range, each word was frequency shifted up by one octave using the STRAIGHT vocoder [29,30]. Each word was adjusted so that the peak intensity (loudest 100 ms of the vowel) was normalized to 60 dB. The words were concatenated to create the word sequences ‘sad wad bad yad’ and ‘sad wad dad yad’ (Figure 1), which will be referred to as “word sequences”. These word sequences were then compressed to 70%, 50%, 30%, 20% and 10% of their original length (2.9, 4, 6.7, 10, and 20 syllables per second, sps) using the STRAIGHT vocoder. Conversational syllable rates in English are estimated to be between 3.5 – 7 sps [31–33]. The words ‘bad’ and ‘dad’ were also presented in isolation at 100% of their original length (500 ms) and temporally compressed to 50% and 30% of their original length (250 and 150 ms).

Vagus nerve stimulation – sound pairing

Vagus nerve stimulation was paired with tone train presentation approximately 300 times per day for 20 days, as in previous studies [17–20]. The VNS + 5 pps paired group experienced an average number of stimulations per day of 314.5 ± 3.5 , while the VNS + 15 pps paired group experienced 314.8 ± 2.0 stimulations per day, which was not significantly different ($p = 0.94$). The onset of VNS was simultaneous with the onset of the third tone in the train of six tones. The VNS was a 500 ms pulse train at 30 Hz with an intensity of 0.8 mA and a biphasic pulse width of 100 μ s. The average interval between sounds was 30 seconds, with an average session length of 2.5 hours. Rats were awake and unrestrained during the pairing sessions that took place in a 25 x 25 x 25 cm³ wire cage. The tone trains were delivered free-field from an Optimus Bullet Horn Tweeter speaker positioned 20 cm above the cage.

Electrophysiology recordings

Following 20 days of VNS sound pairing, primary auditory cortex multiunit responses were recorded, as in previous studies [17,25,34]. Responses were recorded from 262 A1 sites in seven VNS + 15 pps paired rats, 157 A1 sites in four VNS + 5 pps paired rats, and 170 A1 sites in six experimentally naïve control rats. Recording sites were chosen to evenly sample A1 while avoiding blood vessels. There was no significant difference in the number of A1 sites between each of the experimental groups ($F(2,14) = 0.89$, $p = 0.43$). There was no significant difference in the sampled characteristic frequency ranges between each of the experimental groups ($F(2,14) = 2.69$, $p = 0.10$). In addition, there was no significant difference in the recording site density between each of the experimental groups ($F(2,14) = 0.76$, $p = 0.49$). Rats were anesthetized with pentobarbital using the same protocol as the vagus nerve surgery. A tracheotomy was performed and a humidified air tube was provided in order to facilitate breathing. A cisternal drain was

opened in order to reduce brain swelling. A craniotomy and durotomy were performed over right auditory cortex, and 4 Parylene-coated tungsten microelectrodes (FHC Inc., 1 – 2 M Ω impedance) were simultaneously lowered to layer IV/V of primary auditory cortex. At each recording site, tones were randomly interleaved and presented, ranging in frequency from 1 – 48 kHz in 0.0625 octave steps and ranging in intensity from 0 – 75 dB in 5 dB steps. Following tone presentation, tone trains were presented at 70 dB with 2 seconds of silence between each train. Tone trains were randomly interleaved and presented at 11 repetition rates (3, 5, 7, 9, 10, 11, 13, 15, 17, 20, and 25 pps). The carrier frequency of the tones was the paired carrier frequency that was the closest to the characteristic frequency of the recording site. Following tone train presentation, speech sounds were presented both in isolation ('bad' or 'dad') and as word sequences ('sad wad bad yad' or 'sad wad dad yad'), at varying presentation rates from 2 sps to 20 sps.

Data analysis

The response strength evoked by speech sounds was quantified in each group using both the total number of driven spikes as well as the peak firing rate. Response latency was quantified as the peak firing latency. Neural discrimination accuracy was quantified using a nearest-neighbor classifier [25,27,35,36]. The classifier was provided with the 40 ms onset response evoked by the speech sound 'bad', which was compared to the onset response evoked by the speech sound 'dad'. At each recording site, a template of the response pattern evoked by 'bad' and 'dad' was generated from 19 of the 20 repeats recorded for each sound. The response pattern evoked during the remaining single trial response was compared to the average response templates for 'bad' and 'dad'. Spike arrival times were binned with 1 ms precision. Euclidean distance was used to quantify the similarity between the single trial response and the 2 sound

response templates. The single trial response was assigned to the more similar sound response template (minimum Euclidean distance). Bonferonni correction was used to correct for multiple comparisons. Error estimates are standard error of the mean across recording sites. All data analysis was performed using custom MATLAB software.

Results

VNS pairing altered the auditory cortex response to isolated speech sounds

VNS paired with tone trains significantly altered the A1 peak firing rate and peak latency response to speech sounds. Following 20 days of VNS-tone train pairing, there was no significant difference in the total number of driven spikes evoked by speech sounds in either of the experimental groups ($p = 0.11$ for control vs. VNS + 5 pps paired, $p = 0.92$ for control vs. VNS + 15 pps paired, $p = 0.09$ for VNS + 5 pps paired vs. VNS + 15 pps paired, number of spikes evoked in the 600 ms neural response, Figure 2a). While the total number of driven spikes did not change, the peak firing rate in response to speech sounds in the VNS + 15 pps paired group significantly increased by 8% compared to the control group ($p = 0.04$) and increased by 12% compared to the VNS + 5 pps paired group ($p = 0.008$, Figure 2b). This increase in the peak firing rate in the VNS + 15 pps paired group was accompanied by a significant 2% decrease in the peak firing latency compared to the control group ($p = 0.008$), and a 2% decrease in the latency compared to the VNS + 5 pps paired group ($p = 0.02$, Figure 2c). The VNS + 5 pps paired group did not exhibit an alteration in the peak firing rate or latency compared to the control group ($p = 0.4$ for peak firing rate, $p = 0.92$ for peak latency, Figure 2b,c).

VNS pairing with 5 pps or 15 pps tone trains has been previously shown to respectively decrease or increase the temporal following rate of A1 neurons [19]. We hypothesized that pairing VNS with 15 pps trains would improve the A1 response to rapidly presented speech sounds. Although the number of driven spikes evoked by full-length speech sounds (500 ms) was unaltered following VNS-tone train pairing, the number of spikes evoked by compressed speech sounds was significantly increased by 11% in the VNS + 15 pps paired group compared to the control group ($p = 0.04$ for 150 ms sound length, Figures 3 & 4a). There was no difference in the number of driven spikes evoked by compressed sounds in the VNS + 5 pps paired group compared to the VNS + 15 pps paired group or the control group ($p > 0.05$, Figure 4a). The peak firing rate was 7% decreased in the VNS + 5 pps paired group compared to the control group for the 150 ms sound length ($p = 0.04$, Figure 4b), with no other significant differences in the peak firing rate evoked by compressed speech sounds between the groups ($p > 0.05$). The VNS + 15 pps paired group exhibited a significant decrease in the peak firing latency compared to the control group across all presentation rates, indicative of faster neural responses to a speech sound of any duration (16% faster, $p = 0.0003$ for 250 ms sound length; 7% faster, $p = 0.007$ for 150 ms sound length, Figure 4c). The VNS + 15 pps paired group also exhibited a significant decrease in the peak firing latency compared to the VNS + 5 pps paired group across all presentation rates (13% faster, $p = 0.003$ for 250 ms sound length; 7% faster, $p = 0.01$ for 150 ms sound length, Figure 4c). The VNS + 5 pps paired group did not exhibit an alteration in the peak latency compared to the control group at any presentation rates ($p > 0.05$, Figure 4c).

VNS pairing altered the auditory cortex response to word sequences

We next tested the hypothesis that altering the temporal following rate of A1 neurons would affect the ability of these neurons to process sentence-like word sequences (Figure 1c,d).

The word sequences evoked A1 activity in response to each of the speech sounds in the sequence (Figure 5). There was no significant difference in the total number of driven spikes across the duration of the sound between the control group (38.5 ± 2.0 spikes in 2150 ms) and the VNS + 15 pps paired group (39.3 ± 1.8 spikes, $p = 0.77$) or the VNS + 5 pps paired group (35.5 ± 2.2 spikes, $p = 0.31$). There was also no significant difference in the total number of driven spikes across the duration of the sound between the two VNS paired groups ($p = 0.18$). However, the number of driven spikes evoked in the onset response (first 40 ms after presentation of each of the 4 full-length speech sounds in the sequence) was significantly increased by 24 – 56% in the VNS + 15 pps paired group compared to the control group for 3 of the 4 sounds ($p < 0.0063$, Figure 6a). This change was accompanied by a significantly 28 – 59% larger peak firing rate in the VNS + 15 pps paired group compared to the control group for 3 of the 4 sounds ($p < 0.0063$, Figure 6b). Additionally, the peak firing latency was significantly faster by 10% in the VNS + 15 pps paired group compared to the control group ($p = 0.0002$, Figure 6c). The peak firing rate was also significantly increased by 28 – 35% in the VNS + 15 pps paired group compared to the VNS + 5 pps paired group for 3 of the 4 sounds ($p < 0.0063$, Figure 6b). There was no significant difference in the number of driven spikes, peak firing rate, or peak firing latency evoked by the speech sounds in the word sequence between the VNS + 5 pps paired group and the control group ($p > 0.0063$, Figure 6).

In addition to presenting word sequences at 2 sps, which is slower than conversational speech rates, each of the sequences was presented at multiple presentation rates, up to a maximum rate of 20 sps (Figure 7). Across all presentation rates, there was no significant difference in the total number of driven spikes evoked in response to the word sequences across the experimental groups ($F(2,3526) = 2.47$, $p = 0.09$, two-way ANOVA with Bonferroni

correction, Figure 8). As the word sequence presentation rate increased, each of the experimental and control groups exhibited a decrease in the total number of evoked spikes ($F(5,3526) = 126.14$, $p < 0.00001$, two-way ANOVA with Bonferroni correction, Figure 8).

While the words in the first, second, and fourth positions of the word sequence remained constant, two word choices were presented in the third position in order to determine whether VNS-tone train pairing has an effect on the discriminability of the words ‘bad’ and ‘dad’. Similar to the entire word sequence, the number of driven spikes evoked by the onset of ‘bad’ and ‘dad’ was significantly different between the three groups ($F(2,3526) = 18.27$, $p < 0.00001$, Figure 9a). Across all presentation rates for the word sequences, the average number of driven spikes evoked by the third word was 8% larger in the VNS + 15 pps paired group compared to the control group (3.1 ± 0.04 spikes vs. 2.9 ± 0.05 spikes) and 15% larger compared to the VNS + 5 pps paired group (2.7 ± 0.05 spikes). There was no significant change in the number of driven spikes between the VNS + 5 pps paired group and the control group. In addition, the peak firing rate in response to ‘bad’ and ‘dad’ presented as part of a word sequence was significantly different between the three groups ($F(2,3526) = 18.63$, $p < 0.00001$, Figure 9b). Across all presentation rates, the average peak firing rate was 10% weaker in the VNS + 5 pps paired group compared to the control group (474 ± 7 Hz vs. 521 ± 7 Hz) and 12% weaker compared to the VNS + 15 pps paired group (529 ± 6 Hz). There was no significant change in the peak firing rate between the VNS + 15 pps paired group and the control group. The peak firing latency in response to ‘bad’ and ‘dad’ presented as part of a word sequence was also significantly different between the three groups ($F(2,3526) = 23.65$, $p < 0.00001$, Figure 9c). Across all presentation rates, the average peak firing latency was 6% faster in the VNS + 15 pps paired group (28.9 ± 0.2 ms) and 2% faster in the VNS + 5 pps paired group (30.0 ± 0.2 ms) compared to the control group (30.7 ± 0.2

ms). The peak latency was 4% faster in the VNS + 15 pps paired group compared to the VNS + 5 pps paired group.

VNS pairing altered the neural discriminability of speech sounds

A neural classifier was used to determine whether the temporal plasticity generated by pairing VNS with tone trains can improve neural discrimination of isolated words or word sequences. When the words ‘bad’ and ‘dad’ were presented in isolation at a length of 500 ms (100% sound length), the neural classifier was able to accurately discriminate between the A1 response patterns evoked by the two sounds equally in the three groups of rats ($85 \pm 1\%$ correct control; $85 \pm 1\%$ correct VNS + 5 pps; $84 \pm 1\%$ correct VNS + 15 pps; $p > 0.05$, Figure 10a). In the control rats, increased presentation rate decreased the accuracy of the neural classifier by 8% from 85% accurate for full-length sounds to 79% accurate for sounds compressed to 150 ms, and there was no significant difference in neural classifier accuracy between the groups at any presentation rate ($p > 0.05$, Figure 10a).

The alterations in response strength and latency enhanced the ability of a neural classifier to discriminate between a pair of speech sounds in a word sequence. When the words ‘bad’ and ‘dad’ were presented as part of a word sequence, the neural classifier discrimination accuracy was significantly altered between the three groups ($F(2,3526) = 11.51$, $p = 0.00001$, Figure 10b). Across all presentation rates, the neural classification accuracy of ‘bad’ and ‘dad’ was $72.3 \pm 0.4\%$ in the control group, $73.5 \pm 0.5\%$ in the VNS + 5 pps paired group, and $75 \pm 0.4\%$ in the VNS + 15 pps paired group. For example, the neural classification of ‘bad’ and ‘dad’ presented at 6.7 sps was $65 \pm 1\%$ accurate in the control group, and increased by 11% to $72 \pm 0.8\%$ accurate in the VNS + 15 pps paired group ($p < 0.00001$). At the same rate, neural classification

accuracy was 6% larger in the VNS + 15 pps paired group compared to the VNS + 5 pps paired group ($p = 0.005$), suggesting that the stimulus rate paired with VNS is an important parameter.

Discussion

Previous studies have documented that VNS-sound pairing alters the neural responses evoked by the paired sound. This study extends previous findings by documenting that VNS paired with tone trains alters the neural responses evoked by unpaired novel speech sounds. Responses to speech sounds were stronger and faster in the group that received VNS paired with 15 pps tone trains compared to the control group. In addition, a neural classifier was more accurate at discriminating between speech sounds in the VNS + 15 pps paired group. It has recently been shown that VNS paired with upper limb rehabilitation therapy yields three times the benefit compared to rehabilitation therapy alone [37]. The robust changes in neural responses generated by VNS-sound pairing suggest that pairing VNS with auditory training may similarly enhance speech therapy.

A number of studies suggest that underlying neural temporal processing deficits are responsible for speech processing deficits in learning impaired populations [2,38,39]. Children with speech, language, and reading disorders are significantly impaired at discriminating auditory stimuli at rapid rates of presentation [40–42]. For example, individuals with dyslexia have significantly weaker EEG responses when presented with rapid stimuli compared to matched control subjects [43]. In addition, the auditory brainstem and cortex responses in poor readers exhibit significantly more paired pulse depression and increased neural variability compared to good readers [44,45]. These neural deficits have also been documented in rat models of dyslexia. Rats with knockdown of the dyslexia-associated gene *Kiaa0319* exhibit both

weaker auditory cortex responses to speech sounds and significantly degraded A1 responses to rapidly presented noise burst trains [46]. Additionally, rats with knockdown of the dyslexia-associated gene *Dcdc2* exhibit a greatly impaired ability to identify a target speech sound embedded in a stream of non-target speech sounds, despite normal discrimination of isolated speech sounds [26]. In both humans with language disorders and animal models, neural and behavioral temporal processing deficits are commonly observed.

Intensive behavioral training with rapid sounds has been shown to improve both perceptual ability and neural responses [3,13,47–49]. For example, following weeks of auditory training, children with autism exhibited improvements in both brainstem and cortical response timing [48]. This finding is paralleled in rodents; rats that were trained to use the repetition rate of noise pulses in order to find a target location exhibited enhanced A1 responses to rapid stimuli as well as stronger phase-locked responses [7]. Auditory training also greatly benefits rat models of disorders, such as autism or dyslexia [26,50]. Rats prenatally exposed to valproic acid are a well-studied autism model that exhibit significantly weaker and less synchronized responses to rapidly presented stimuli [51]. Months of speech discrimination training normalizes both the degree of synchronization and the strength of the response to rapid sounds [52].

Similar to intense behavioral training, numerous studies have documented that VNS pairing therapy alters both behavioral performance as well as cortical responses. Although translation from animal models to clinical effectiveness is notoriously difficult, recent studies suggest that VNS pairing therapy may translate from animal studies to clinical benefit. For example, VNS paired with sound therapy has been shown to improve tinnitus symptoms in both tinnitus patients [53,54] and an animal model of tinnitus [17]. Similarly, VNS paired with motor therapy has been shown to improve upper limb function and neural responses in stroke patients

[37] and multiple animal models of stroke [23,55,56]. VNS paired with 15 pps tone trains increases the ability of A1 neurons to respond to rapid sounds, while VNS paired with 5 pps tone trains decreases the ability of A1 neurons to respond to rapid sounds [19]. The current study documented that VNS paired with 15 pps tone trains can additionally improve neural responses to unpaired novel speech sounds. In particular, neural discrimination accuracy was improved at sound presentation rates between 4 sps and 10 sps, which encompasses a presentation rate range considered to be ‘conversational’ speech [31–33,42]. Future studies are needed to determine whether pairing VNS with speech therapy could potentially be used to improve speech perception and production as well as neural responses to speech sounds in individuals with receptive language disorders.

In particular, future experiments are necessary to determine which VNS pairing paradigm would best improve speech processing for specific subtypes of communication disorders. For example, animals and individuals with tinnitus experienced VNS paired with tones of varying frequencies excluding the tinnitus frequency in order to shrink the auditory cortex representation of the tinnitus frequency and expand and normalize the representation of all other frequencies [17,53,54]. In order to enhance speech processing, individuals that exhibit weak auditory cortex responses to sound [57–59] may most benefit from strengthening the auditory cortex response to sounds, while individuals that exhibit rapid temporal processing impairments [40,41] may most benefit from improving the auditory cortex response to rapid sounds. Non-speech sounds could potentially be paired with VNS in order to improve specific aspects of speech processing. For example, rapid stimuli could be paired with VNS to improve problems with distinguishing rapid temporal transitions found in consonants [19,60], such as voice onset time distinctions or place of articulation distinctions. Similarly, ripple sounds could be paired with VNS to improve

problems with distinguishing spectral transitions found in vowels [61], while frequency modulated sweeps could be paired with VNS to accelerate the learning of tonal languages [62–64].

There are limitations of the current study that should be explored in follow up studies. For example, it is possible that the pentobarbital anesthesia masked the VNS pairing induced changes in A1 responses [65]. Previous studies have documented suppressed responses to repetitive stimuli in anesthetized compared to awake auditory cortex recordings. For example, one study found that responses to the first click in a train were not significantly different between awake and anesthetized recording states, but responses to subsequent clicks occurred later and were weaker in anesthetized recordings compared to awake recordings [66]. For the current study, our analysis used only the initial onset response to each speech sound, which is unlikely to be significantly affected. Additionally, our previous study documented that anesthetized A1 responses to speech sounds can be used to predict behavioral speech discrimination ability just as accurately as A1 responses to speech sounds recorded from awake passively listening animals [25].

Many previous studies revealed large behavioral and neurophysiological differences when VNS was paired with sensory or motor events, but no behavioral or neurophysiological differences in rats that received VNS alone or in rats that were exposed to sensory or motor events without VNS pairing [17,19,21–23]. The control rats in the current study did not experience exposure to sounds or exposure to VNS. It is possible that sound exposure alone or VNS alone could have altered A1 responses. The temporal specificity of the plasticity documented in this study and the previous literature suggests that the changes documented here were not likely to result from passive exposure to sounds or VNS.

Conclusions

Pairing VNS with 15 pps tone trains increases the following rate of A1 neurons and results in stronger, faster, and more discriminable responses to speech sounds. Future preclinical and clinical studies are needed to evaluate the potential of VNS rehabilitation for communication disorders, such as aphasia and autism.

Acknowledgements

The authors would like to thank K. Ram and R. Cheung for their assistance with microelectrode recordings. We would like to thank S. Sudanagunta, M. Fink, H. Rasul, D. Vuppalla, A. Kuzu, C. Walker, C. Omana and M. Borland for their help with daily VNS pairing sessions, and T. Centanni, S. Hays, and J. Riley for helpful discussions on earlier versions of the manuscript. This research was supported by grants from the National Institutes of Health to MPK (Grant # R01DC010433 & R44DC010084). This program was supported by the Defense Advanced Research Projects Agency (DARPA) Biological Technologies Office (BTO) Electrical Prescriptions (ElectRx) program under the auspices of Dr. Doug Weber through the Space and Naval Warfare Systems Center, Pacific Cooperative Agreement No. HR0011-15-2-0017 and N66001-15-2-4057. The funding sources had no role in study design, the collection, analysis and interpretation of data, writing of the report, or the decision to submit the article for publication.

Conflict of interest statement

NDE is an employee of MicroTransponder Inc. and MPK is a paid consultant for MicroTransponder Inc.

References

- [1] Ahissar E, Nagarajan S, Ahissar M, Protopapas A, Mahncke H, Merzenich MM. Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proc Natl Acad Sci U S A* 2001;98:13367–72.
- [2] Oram Cardy JE, Flagg EJ, Roberts W, Brian J, Roberts TPL. Magnetoencephalography identifies rapid temporal processing deficit in autism and language impairment. *Neuroreport* 2005;16:329–32.
- [3] Merzenich MM, Jenkins WM, Johnston P, Schreiner C, Miller SL, Tallal P. Temporal processing deficits of language-learning impaired children ameliorated by training. *Science* 1996;271:77–81.
- [4] Tallal P. Auditory temporal perception, phonics, and reading disabilities in children. *Brain Lang* 1980;9:182–98.
- [5] Robson H, Grube M, Lambon Ralph MA, Griffiths TD, Sage K. Fundamental deficits of auditory perception in Wernicke’s aphasia. *Cortex* 2013;49:1808–22.
- [6] Talvitie SS, Matilainen LE, Pekkonen E, Alku P, May PJC, Tiitinen H. The effects of cortical ischemic stroke on auditory processing in humans as indexed by transient brain responses. *Clin Neurophysiol* 2010;121:912–20.
- [7] Bao S, Chang EF, Woods J, Merzenich MM. Temporal plasticity in the primary auditory cortex induced by operant perceptual learning. *Nat Neurosci* 2004;7:974–81.
- [8] Zhou X, Merzenich MM. Developmentally degraded cortical temporal processing restored by training. *Nat Neurosci* 2009;12:26–8.
- [9] Beitel RE, Vollmer M, Raggio MW, Schreiner CE. Behavioral training enhances cortical temporal processing in neonatally deafened juvenile cats. *J Neurophysiol* 2011;106:944–59.
- [10] Vollmer M, Beitel RE. Behavioral training restores temporal processing in auditory cortex of long-deaf cats. *J Neurophysiol* 2011;106:2423–36.
- [11] Zhu X, Liu X, Wei F, Wang F, Merzenich MM, Schreiner CE, et al. Perceptual Training Restores Impaired Cortical Temporal Processing Due to Lead Exposure. *Cereb Cortex* 2016;26:334–45.
- [12] Rojas Sosa MC, Fraire Martínez MI, Olvera Gómez JL, Jáuregui-Renaud K. Early auditory middle latency evoked potentials correlates with recovery from aphasia after stroke. *Clin Neurophysiol* 2009;120:136–9.
- [13] Becker F, Reinvang I. Event-related potentials indicate bi-hemispherical changes in speech sound processing during aphasia rehabilitation. *J Rehabil Med* 2007;39:658–61.
- [14] Tallal P, Miller S, Jenkins W, Merzenich M. The role of temporal processing in developmental language-based learning disorders: Research and clinical implications. *Found. Read. Acquis. dyslexia Implic. early Interv.*, 1997, p. 49–66.
- [15] Gillam RB, Loeb DF, Hoffman LM, Bohman T, Champlin CA, Thibodeau L, et al. The

- 425 efficacy of Fast ForWord Language intervention in school-age children with language
426 impairment: a randomized controlled trial. *J Speech Lang Hear Res* 2008;51:97–119.
- 427 [16] Kilgard MP, Merzenich MM. Cortical Map Reorganization Enabled by Nucleus Basalis
428 Activity. *Science* (80-) 1998;279:1714–8.
- 429 [17] Engineer ND, Riley JR, Seale JD, Vrana WA, Shetake JA, Sudanagunta SP, et al.
430 Reversing pathological neural activity using targeted plasticity. *Nature* 2011;470:101–4.
- 431 [18] Borland MS, Vrana WA, Moreno NA, Fogarty EA, Buell EP, Sharma P, et al. Cortical
432 Map Plasticity as a Function of Vagus Nerve Stimulation Intensity. *Brain Stimul*
433 2015;9:117–23.
- 434 [19] Shetake JA, Engineer ND, Vrana WA, Wolf JT, Kilgard MP. Pairing tone trains with
435 vagus nerve stimulation induces temporal plasticity in auditory cortex. *Exp Neurol*
436 2012;233:342–9.
- 437 [20] Engineer CT, Engineer ND, Riley JR, Seale JD, Kilgard MP. Pairing Speech Sounds
438 With Vagus Nerve Stimulation Drives Stimulus-specific Cortical Plasticity. *Brain Stimul*
439 2015;8:637–44.
- 440 [21] Porter BA, Khodaparast N, Fayyaz T, Cheung RJ, Ahmed SS, Vrana WA, et al.
441 Repeatedly pairing vagus nerve stimulation with a movement reorganizes primary motor
442 cortex. *Cereb Cortex* 2012;22:2365–74.
- 443 [22] Hays SA, Khodaparast N, Ruiz A, Sloan AM, Hulsey DR, Rennaker RL, et al. The timing
444 and amount of vagus nerve stimulation during rehabilitative training affect poststroke
445 recovery of forelimb strength. *Neuroreport* 2014;25:676–82.
- 446 [23] Khodaparast N, Hays SA, Sloan AM, Fayyaz T, Hulsey DR, Rennaker RL, et al. Vagus
447 Nerve Stimulation Delivered During Motor Rehabilitation Improves Recovery in a Rat
448 Model of Stroke. *Neurorehabil Neural Repair* 2014.
- 449 [24] Kilgard MP, Merzenich MM. Plasticity of temporal information processing in the primary
450 auditory cortex. *Nat Neurosci* 1998;1:727–31.
- 451 [25] Engineer CT, Perez CA, Chen YH, Carraway RS, Reed AC, Shetake JA, et al. Cortical
452 activity patterns predict speech discrimination ability. *Nat Neurosci* 2008;11:603–8.
- 453 [26] Centanni TM, Booker AB, Chen F, Sloan AM, Carraway RS, Rennaker RL, et al.
454 Knockdown of Dyslexia-Gene *Dcdc2* Interferes with Speech Sound Discrimination in
455 Continuous Streams. *J Neurosci* 2016;36:4895–906.
- 456 [27] Perez CA, Engineer CT, Jakkamsetti V, Carraway RS, Perry MS, Kilgard MP. Different
457 timescales for the neural coding of consonant and vowel sounds. *Cereb Cortex*
458 2013;23:670–83.
- 459 [28] Shetake JA, Wolf JT, Cheung RJ, Engineer CT, Ram SK, Kilgard MP. Cortical activity
460 patterns predict robust speech discrimination ability in noise. *Eur J Neurosci*
461 2011;34:1823–38.
- 462 [29] Kawahara H. Speech representation and transformation using adaptive interpolation of

- 463 weighted spectrum: Vocoder revisited. *Proc ICASSP* 1997;2:1303–6.
- 464 [30] Kelly JB, Masterton B. Auditory sensitivity of the albino rat. *J Comp Physiol Psychol*
465 1977;91:930–6.
- 466 [31] Pellegrino F, Coupé C, Marsico E. Across-Language Perspective on Speech Information
467 Rate. *Language (Baltim)* 2011;87:539–58.
- 468 [32] Goldman-Eisler F. The significance of changes in the rate of articulation. *Lang Speech*
469 1961;4:171–4.
- 470 [33] Jacewicz E, Fox RA, Wei L. Between-speaker and within-speaker variation in speech
471 tempo of American English. *J Acoust Soc Am* 2010;128:839–50.
- 472 [34] Polley DB, Read HL, Storace DA, Merzenich MM. Multiparametric auditory receptive
473 field organization across five cortical fields in the albino rat. *J Neurophysiol*
474 2007;97:3621–38.
- 475 [35] Centanni TM, Engineer CT, Kilgard MP. Cortical speech-evoked response patterns in
476 multiple auditory fields are correlated with behavioral discrimination ability. *J*
477 *Neurophysiol* 2013;110:177–89.
- 478 [36] Foffani G, Moxon KA. PSTH-based classification of sensory stimuli using ensembles of
479 single neurons. *J Neurosci Methods* 2004;135:107–20.
- 480 [37] Dawson J, Pierce D, Dixit A, Kimberley TJ, Robertson M, Tarver B, et al. Safety,
481 Feasibility, and Efficacy of Vagus Nerve Stimulation Paired With Upper-Limb
482 Rehabilitation After Ischemic Stroke. *Stroke* 2016;47:143–50.
- 483 [38] Auther LL, Wertz RT, Miller T a., Kirshner HS. Relationships among the mismatch
484 negativity (MMN) response, auditory comprehension, and site of lesion in aphasic adults.
485 *Aphasiology* 2000;14:461–70.
- 486 [39] O'Connor K. Auditory processing in autism spectrum disorder: a review. *Neurosci*
487 *Biobehav Rev* 2012;36:836–54.
- 488 [40] Tallal P, Miller S, Fitch RH. Neurobiological basis of speech: a case for the preeminence
489 of temporal processing. *Ann N Y Acad Sci* 1993;682:27–47.
- 490 [41] Tallal P, Piercy M. Defects of non-verbal auditory perception in children with
491 developmental aphasia. *Nature* 1973;241:468–9.
- 492 [42] Abrams DA, Nicol T, Zecker S, Kraus N. Abnormal cortical processing of the syllable
493 rate of speech in poor readers. *J Neurosci* 2009;29:7686–93.
- 494 [43] McAnally KI, Stein JF. Scalp potentials evoked by amplitude-modulated tones in
495 dyslexia. *J Speech Lang Hear Res* 1997;40:939–45.
- 496 [44] Nagarajan S, Mahncke H, Salz T, Tallal P, Roberts T, Merzenich MM. Cortical auditory
497 signal processing in poor readers. *Proc Natl Acad Sci U S A* 1999;96:6483–8.
- 498 [45] Hornickel J, Kraus N. Unstable representation of sound: A biological marker of dyslexia.
499 *J Neurosci* 2013;33:3500–4.

- 500 [46] Centanni TM, Booker AB, Sloan AM, Chen F, Maher BJ, Carraway RS, et al.
 501 Knockdown of the Dyslexia-Associated Gene *Kiaa0319* Impairs Temporal Responses to
 502 Speech Stimuli in Rat Primary Auditory Cortex. *Cereb Cortex* 2014;24:1753–66.
- 503 [47] Tallal P, Miller SL, Bedi G, Byma G, Wang X, Nagarajan SS, et al. Language
 504 comprehension in language-learning impaired children improved with acoustically
 505 modified speech. *Science* (80-) 1996;271:81–4.
- 506 [48] Russo NM, Hornickel J, Nicol T, Zecker S, Kraus N. Biological changes in auditory
 507 function following training in children with autism spectrum disorders. *Behav Brain Funct*
 508 2010;6:60.
- 509 [49] Warrier CM, Johnson KL, Hayes EA, Nicol T, Kraus N. Learning impaired children
 510 exhibit timing deficits and training-related improvements in auditory cortical responses to
 511 speech in noise. *Exp Brain Res* 2004;157:431–41.
- 512 [50] Engineer CT, Rahebi KC, Borland MS, Buell EP, Centanni TM, Fink MK, et al.
 513 Degraded neural and behavioral processing of speech sounds in a rat model of Rett
 514 syndrome. *Neurobiol Dis* 2015;83:26–34.
- 515 [51] Engineer CT, Centanni TM, Im KW, Borland MS, Moreno NA, Carraway RS, et al.
 516 Degraded auditory processing in a rat model of autism limits the speech representation in
 517 non-primary auditory cortex. *Dev Neurobiol* 2014;74:972–86.
- 518 [52] Engineer CT, Centanni TM, Im KW, Kilgard MP. Speech sound discrimination training
 519 improves auditory cortex responses in a rat model of autism. *Front Syst Neurosci*
 520 2014;8:137.
- 521 [53] De Ridder D, Vanneste S, Engineer ND, Kilgard MP. Safety and efficacy of vagus nerve
 522 stimulation paired with tones for the treatment of tinnitus: a case series. *Neuromodulation*
 523 2014;17:170–9.
- 524 [54] De Ridder D, Kilgard M, Engineer N, Vanneste S. Placebo-controlled vagus nerve
 525 stimulation paired with tones in a patient with refractory tinnitus: a case report. *Otol*
 526 *Neurotol* 2015;36:575–80.
- 527 [55] Hays SA, Khodaparast N, Hulsey DR, Ruiz A, Sloan AM, Rennaker RL, et al. Vagus
 528 nerve stimulation during rehabilitative training improves functional recovery after
 529 intracerebral hemorrhage. *Stroke* 2014;45:3097–100.
- 530 [56] Khodaparast N, Hays SA, Sloan AM, Hulsey DR, Ruiz A, Pantoja M, et al. Vagus nerve
 531 stimulation during rehabilitative training improves forelimb strength following ischemic
 532 stroke. *Neurobiol Dis* 2013;60:80–8.
- 533 [57] Naito Y, Hirano S, Honjo I, Okazawa H, Ishizu K, Takahashi H, et al. Sound-induced
 534 activation of auditory cortices in cochlear implant users with post- and prelingual deafness
 535 demonstrated by positron emission tomography. *Acta Otolaryngol* 1997;117:490–6.
- 536 [58] Sharma A, Dorman MF, Spahr AJ. A sensitive period for the development of the central
 537 auditory system in children with cochlear implants: implications for age of implantation.
 538 *Ear Hear* 2002;23:532–9.

- 539 [59] Stach BA, Stoner WR, Smith SL, Jerger JF. Auditory evoked potentials in Rett syndrome.
540 J Am Acad Audiol 1994;5:226–30.
- 541 [60] Kilgard MP, Merzenich MM. Plasticity of temporal information processing in the primary
542 auditory cortex. Nat Neurosci 1998;1:727–31.
- 543 [61] Kilgard MP, Pandya PK, Vazquez JL, Rathbun DL, Engineer ND, Moucha R. Spectral
544 features control temporal plasticity in auditory cortex. Audiol Neurotol 2001;6:196–202.
- 545 [62] Moucha R, Pandya PK, Engineer ND, Rathbun DL, Kilgard MP. Background sounds
546 contribute to spectrotemporal plasticity in primary auditory cortex. Exp Brain Res
547 2005;162:417–27.
- 548 [63] Wang Y, Spence MM, Jongman A, Sereno J a. Training American listeners to perceive
549 Mandarin tones. J Acoust Soc Am 1999;106:3649–58.
- 550 [64] Wang Y, Sereno JA, Jongman A, Hirsch J. fMRI evidence for cortical modification
551 during learning of Mandarin lexical tone. J Cogn Neurosci 2003;15:1019–27.
- 552 [65] Cheung SW, Nagarajan SS, Bedenbaugh PH, Schreiner CE, Wang X, Wong A. Auditory
553 cortical neuron response differences under isoflurane versus pentobarbital anesthesia.
554 Hear Res 2001;156:115–27.
- 555 [66] Rennaker RL, Carey HL, Anderson SE, Sloan AM, Kilgard MP. Anesthesia suppresses
556 nonsynchronous responses to repetitive broadband stimuli. Neuroscience 2007;145:357–
557 69.
- 558
- 559
- 560

Figure legends

Figure 1. Spectrograms of the speech sounds used in this study. The isolated speech words **(a)** bad and **(b)** dad were 500 ms long. Time is represented on the x axis and frequency is represented on the y axis. The word sequences **(c)** ‘sad wad bad yad’ and **(d)** ‘sad wad dad yad’ were 2.2 seconds long.

Figure 2. VNS-tone train pairing altered the auditory cortex response to isolated speech sounds. **(a)** The number of driven spikes evoked across recording sites in response to isolated speech sounds was equivalent in VNS-tone train paired rats compared to control rats. The number of driven spikes was quantified using the 600 ms A1 response to each sound. Error bars indicate SEM across recording sites. **(b)** The peak firing rate was strengthened in the VNS + 15 pps paired group compared to the control group. Stars indicate a statistically significant difference between the VNS + 15 pps paired group and the control group ($p < 0.05$). **(c)** The peak firing latency was significantly shorter in the VNS + 15 pps paired group compared to the control group.

Figure 3. Primary auditory cortex post-stimulus time histogram (PSTH) response to the isolated speech sound ‘dad’ at varying presentation rates. The sound ‘dad’ was presented at **(a)** 500 ms, **(b)** 250 ms, and **(c)** 150 ms. The waveform for the ‘dad’ sounds are plotted in gray above the PSTH responses.

Figure 4. VNS-tone train pairing altered the auditory cortex response to isolated speech sounds at varying presentation rates. **(a)** The number of driven spikes in response to rapidly presented speech sounds was increased in the VNS + 15 pps paired group compared to the control group. Error bars indicate SEM across recording sites. Green stars indicate a statistically significant difference between the VNS + 15 pps paired group and the control group, while red stars indicate a significant difference between the VNS + 5 pps paired group and the control group, and black stars indicate a significant difference between the VNS + 5 pps paired group and the VNS + 15 pps paired group ($p < 0.05$). **(b)** The peak firing rate was increased in the VNS + 15 pps paired group and decreased in the VNS + 5 pps paired group compared to the control group. **(c)** The peak firing latency was significantly shorter in the VNS + 15 pps paired group compared to the control group at all presentation rates.

Figure 5. The A1 PSTH response to the word sequence ‘sad wad dad yad’.

Figure 6. VNS-tone train pairing altered the auditory cortex response to word sequences. **(a)** The number of driven spikes evoked in response to word sequences was increased in the VNS + 15 pps paired group compared to the control group. Error bars indicate SEM across recording sites. Green stars indicate a statistically significant difference between the VNS + 15 pps paired group and the control group, while black stars indicate a significant difference between the VNS + 5 pps paired group and the VNS + 15 pps paired group (Bonferroni corrected, $p < 0.0063$). **(b)** The peak firing rate in response to word sequences was strengthened in the VNS + 15 pps paired group compared to the control group. **(c)** The peak firing latency was significantly shorter in the

VNS + 15 pps paired group compared to the control group.

Figure 7. The A1 PSTH response to the word sequence ‘sad wad dad yad’ at varying presentation rates. The ‘sad wad bad yad’ and ‘sad wad dad yad’ word sequences were presented at rates from 2 – 20 sps.

Figure 8. The number of driven spikes evoked by word sequences was unaltered following VNS-tone train pairing at multiple presentation rates (Bonferroni corrected, $p > 0.0042$). Error bars indicate SEM across recording sites.

Figure 9. VNS-tone train pairing altered the auditory cortex response to word sequences at varying presentation rates. **(a)** The number of driven spikes evoked by ‘bad’ and ‘dad’ as part of a word sequence was increased in the VNS + 15 pps paired group and decreased in the VNS + 5 pps paired group compared to the control group at certain presentation rates. Error bars indicate SEM across recording sites. Green stars indicate a statistically significant difference between the VNS + 15 pps paired group and the control group, while red stars indicate a significant difference between the VNS + 5 pps paired group and the control group, and black stars indicate a significant difference between the VNS + 5 pps paired group and the VNS + 15 pps paired group (Bonferroni corrected, $p < 0.0042$). **(b)** The peak firing rate in response to ‘bad’ and ‘dad’ was decreased in the VNS + 5 pps paired group compared to the control group. **(c)** The peak firing latency in response to the words ‘dad’ and ‘bad’ was significantly shorter in the VNS + 15 pps paired group compared to the control group across multiple presentation rates.

626

627 **Figure 10.** VNS-tone train pairing altered neural classification accuracy. **(a)** Neural
628 classification accuracy of the isolated speech sounds ‘bad’ vs. ‘dad’ was unaltered following
629 VNS pairing. Error bars indicate SEM across recording sites. The gray shading indicates the
630 range that is considered to be a conversational speech rate. **(b)** Neural classification accuracy of
631 ‘bad’ vs. ‘dad’ when part of a word sequence was enhanced in the VNS + 15 pps paired group
632 compared to control rats ($p = 0.00001$) for conversational rates, but not for faster or slower rates.
633 Green stars indicate a statistically significant difference between the VNS + 15 pps paired group
634 and the control group, while black stars indicate a significant difference between the VNS + 5
635 pps paired group and the VNS + 15 pps paired group.

Figure 1

Figure 2

Figure 3

Figure 4

Figure 5

Figure 6

Figure 7

Figure 8

Figure 9

Figure 10