



OPEN ACCESS

EDITED BY

William Matchin,
University of South Carolina, United States

REVIEWED BY

Jonathan Andrew Berken,
Children's Hospital of Philadelphia,
United States

Ian Phillips,
Walter Reed National Military Medical Center,
United States

*CORRESPONDENCE

Claire T. Honda
✉ claire.honda@mail.mcgill.ca

RECEIVED 18 March 2024

ACCEPTED 08 July 2024

PUBLISHED 25 July 2024

CITATION

Honda CT, Bhutani N, Clayards M and Baum S
(2024) No clear benefit of transcutaneous
auricular vagus nerve stimulation for
non-native speech sound learning.
Front. Lang. Sci. 3:1403080.
doi: 10.3389/flang.2024.1403080

COPYRIGHT

© 2024 Honda, Bhutani, Clayards and Baum.
This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that
the original publication in this journal is cited,
in accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

No clear benefit of transcutaneous auricular vagus nerve stimulation for non-native speech sound learning

Claire T. Honda^{1,2*}, Neha Bhutani³, Meghan Clayards^{2,4,5} and Shari Baum^{2,4}

¹Integrated Program in Neuroscience, McGill University, Montreal, QC, Canada, ²Centre for Research on Brain, Language and Music, Montreal, QC, Canada, ³Revaï, Montreal, QC, Canada, ⁴School of Communication Sciences and Disorders, McGill University, Montreal, QC, Canada, ⁵Department of Linguistics, McGill University, Montreal, QC, Canada

Introduction: Learning to understand and speak a new language can be challenging and discouraging for adults. One potential tool for improving learning is transcutaneous auricular vagus nerve stimulation (taVNS), which modulates perception, memory, and attention systems. It has recently been reported that taVNS can improve English speakers' ability to perceive unfamiliar Mandarin tones. The current project explored the potential benefits of taVNS for language learning beyond tone perception.

Methods: We studied adults' ability to perceive and produce unfamiliar speech sounds as well as any potential change in language learning motivation from pre- to post-training. Forty-five native English speakers were divided into three groups and were trained to perceive German sounds: one group received stimulation during easier-to-learn sounds (vowels), one group received stimulation during harder-to-learn sounds (fricatives), and a control group received no stimulation.

Results and discussion: We did not find evidence that taVNS improved perception or production of the German sounds, but there was evidence that it did improve some aspects of motivation. Specifically, the group that received taVNS during easier sounds showed a significant decrease in feelings of tension/pressure about language learning, while the other groups did not. Overall, the present study does not find that taVNS holds benefits for the acquisition of new speech sounds; however, the field is nascent, and so the potential applications of taVNS for language learning remain to be clarified.

KEYWORDS

transcutaneous auricular vagus nerve stimulation, phonetic perception, non-native perception, speech perception training, phonetic production, language learning motivation

1 Introduction

Many adult learners struggle to attain native-like performance across various measures of linguistic aptitude (e.g., [Munro and Mann, 2005](#); [Abrahamsson and Hyltenstam, 2009](#)). Indeed, language learning outcomes tend to worsen as age of acquisition increases ([Johnson and Newport, 1989](#); [Pulvermüller and Schumann, 1994](#); [Weber-Fox and Neville, 1996](#); [Kang and Guion, 2006](#); [Abrahamsson and Hyltenstam, 2009](#); [White et al., 2013](#)). One aspect of language acquisition that poses particular challenges for adult learners is the perception and production of new speech sounds ([Iverson et al., 2003](#); [Díaz et al., 2012](#)). While infants implicitly learn to differentiate language sounds through unsupervised

exposure (Maye et al., 2002), adults can benefit from explicit instruction in the form of supervised phonetic training paradigms (McCandliss et al., 2002; Iverson et al., 2005); and even after training, many adults still show relatively poor differentiation of new speech sounds in perception and production (Strange and Dittmann, 1984; Hanulíková et al., 2012). The maturation of the brain has been argued to be a contributing factor in these age-related learning differences (see Stowe and Sabourin, 2005, for a review). One potential means of increasing the plasticity of the brain and improving learning in adulthood is through the use of neurostimulation techniques such as vagus nerve stimulation (VNS).

The vagus nerve is the longest cranial nerve in the body, reaching from the medulla down to the colon and innervating multiple organs along the way (Yuan and Silberstein, 2016a). The nerve's afferent fibers send sensory input to the vagal nuclei, which then pass the information along to various brain regions implicated in memory, perception, arousal, and affect, including the locus coeruleus, raphe nucleus, amygdala, thalamus, hippocampus, and nucleus accumbens (Sawchenko, 1983; Berthoud and Neuhuber, 2000; Frangos et al., 2015; Yuan and Silberstein, 2016a). By modifying the activity of the vagus nerve through stimulation, it is therefore possible to alter the activity of multiple brain areas and impact their associated functions (Frangos et al., 2015; Yuan and Silberstein, 2016b)—an approach that has advantages over other methods that only modulate localized neural activity (e.g., transcranial magnetic stimulation, transcranial direct current stimulation, or direct chemical stimulation; Bandler, 1969; Hallett, 2000; Thair et al., 2017). Indeed, in both animals and humans, VNS via an implanted electrode can improve memory (Clark et al., 1995, 1999; Ghacibeh et al., 2006; Sun et al., 2017), likely by modulating synaptic plasticity in the hippocampus (Zuo et al., 2007). Similarly, VNS has been shown to increase arousal and alertness as quantified by behavioral and neural measures (Rizzo et al., 2003; Collins et al., 2021). Such increases in arousal appear to be the result of enhanced excitatory activity in the locus coeruleus and other subcortical structures, leading to widespread activation throughout the cortex (Collins et al., 2021). In addition, VNS can improve positive affect by promoting the release of serotonin and noradrenaline from the raphe nucleus and locus coeruleus (Elger et al., 2000; Austelle et al., 2022). The positive effects of VNS on mood are further attested to by its approved use as a treatment for major depressive disorder (Austelle et al., 2022). In terms of modulating perception, rodent and human studies have also demonstrated that VNS can be paired with auditory stimuli to induce lasting, stimulus-specific plasticity in the auditory cortex (Shetake et al., 2012; De Ridder et al., 2014; Engineer et al., 2015; Lai and David, 2021).

Although VNS shows great promise as a neuromodulatory technique, it involves surgically implanting electrodes in the neck, rendering it invasive and inaccessible to the majority of the population. More recently, transcutaneous auricular vagus nerve stimulation or taVNS—a non-invasive counterpart to VNS—has been introduced as a similarly effective means of modulating neural activity (Frangos et al., 2015; Van Leusden et al., 2015; Yakunina et al., 2017). The auricular branch of the vagus nerve passes just under the skin of the outer ear (cymba concha, cymba cavum, external acoustic meatus, and tragus), and so taVNS can

be administered in a straightforward and accessible way by placing electrodes against the ear (Frangos et al., 2015; Yakunina et al., 2017; Badran et al., 2018; Butt et al., 2020).

The brain regions affected by taVNS are similar to those affected by VNS; these include the locus coeruleus, raphe nucleus, amygdala, insula, thalamus, hippocampus, and nucleus accumbens (Frangos et al., 2015; Yakunina et al., 2017; Badran et al., 2018). As with VNS, taVNS has been found to modulate human perception and cognition. For example, studies have shown that taVNS can improve memory (Jacobs et al., 2015; Sun et al., 2021; Thakkar et al., 2023), arousal (Sharon et al., 2021; Chen et al., 2023), mood (Ferstl et al., 2022), tinnitus symptoms (Shim et al., 2015), and interoception (Villani et al., 2019), as well as decrease reaction times (Chen et al., 2021).

Learning a new language depends crucially on the ability to attend to and remember newly learned information. Thus, taVNS may hold promise for accelerating language learning due to its effects on arousal and memory. Indeed, preliminary evidence suggests that taVNS can enhance memory for spoken and written word lists under some conditions (Giraudier et al., 2020; Kaan et al., 2021), and can enhance reading skills in adults learning a new orthography (Thakkar et al., 2020). The potential for taVNS to aid language learning is also implied by work showing that arousal state modulates phonetic perception (Schuerman et al., 2022). Beyond its effects on arousal and memory, taVNS may enhance language learning by increasing auditory plasticity in the adult brain. As previously mentioned, invasive VNS can lead to stimulus-specific plasticity in the auditory cortex (Shetake et al., 2012; De Ridder et al., 2014; Engineer et al., 2015); and studies of event-related potentials have found that delivering taVNS during auditory perception tasks can enhance auditory preattentive processing (N1 amplitude; Rufener et al., 2023), selective attention (P3 amplitude; Rufener et al., 2018), and lexico-semantic encoding (N400 amplitude; Phillips et al., 2021), as well as decrease auditory processing time (P3 latency; Rufener et al., 2018). Perhaps taVNS therefore induces complementary benefits to arousal, memory, and auditory processing, which work in concert to facilitate the acquisition of a new language.

Important evidence that taVNS may accelerate language learning in adulthood comes from recent work by Llanos et al. (2020). The authors demonstrated that taVNS, in conjunction with perceptual training, can enhance native English speakers' ability to label certain Mandarin tones (Llanos et al., 2020). They divided participants into three groups: one that received stimulation during easier-to-perceive tones, one that received stimulation during harder-to-perceive tones, and a control group that received no stimulation. They found that taVNS specifically enhanced the labeling of easier-to-perceive tones for the group that was stimulated during those tones; this is likely because taVNS can increase arousal, and arousal specifically enhances memory for perceptually salient (i.e., easier-to-perceive) stimuli (Mather and Sutherland, 2011; Mather et al., 2016; Llanos et al., 2020). Nonetheless, a follow-up study by the same research group did not find any overall differences in tone learning performance between stimulated groups and a control group (McHaney et al., 2023). Exploratory analyses did reveal that participants' tone labeling accuracy increased faster for trials during which taVNS

was administered compared to trials without taVNS, and that this effect was most pronounced when taVNS amplitude was low (McHaney et al., 2023). However, these effect sizes were small and there was no main effect of stimulation, so taVNS did not improve overall accuracy (McHaney et al., 2023). Around the same time, Pandža et al. (2020) investigated native English speakers' ability to associate meaning with Mandarin pseudowords that differed in lexical tone. They found that taVNS during Mandarin lexical tone training led to enhanced performance on a subsequent meaning recognition task (Pandža et al., 2020). Additionally, they found that taVNS before training was associated with greater decreases in reaction time on the recognition task and with more accurate performance on a recall task; however, these findings did not hold when taVNS was administered during (rather than before) training. Similar mixed results were obtained by Phillips et al. (2021): taVNS before or during Mandarin lexical tone training did not improve performance on a word learning task, but taVNS before training sped up reaction times on a recognition task, and taVNS during training led to improved recognition of mismatch trials on the recognition task. As such, taVNS appears to have some potential for enhancing lexical tone learning, but effects are not always consistently found; the extent of its efficacy and whether this may generalize beyond the learning of lexical tones remains unclear. Furthermore, while taVNS has shown some benefits for the *perception* of unfamiliar language sounds, it is not yet known whether those benefits extend to the *production* of unfamiliar language sounds.

Apart from facilitating language learning itself, taVNS may improve the subjective learning experience through its effects on mood and motivation. Calloway et al. (2020) found that participants who received taVNS prior to being trained on a language learning task showed greater reductions in negative affect from pre- to post-training compared to control participants who did not receive stimulation. Improvements in mood could subsequently impact motivation and learning outcomes, given that mood is a factor affecting learners' perceptions of success and failure during language learning (Williams et al., 2004). taVNS also seems to play a role in motivation, having been shown to increase adults' motivation to obtain food rewards (Neuser et al., 2020). Yet, to our knowledge, the potential impacts of taVNS on language learning motivation have not been investigated to date.

The current project had the broad aim of determining whether taVNS can enhance the learning of unfamiliar non-tonal speech sound contrasts. More specifically, we had 3 objectives: to determine whether taVNS during non-native speech perception training (1) enhances the *perception* of the trained sounds, (2) enhances the subsequent *production* of the same speech sounds, and (3) enhances *motivation* associated with language learning. The first objective was addressed by running a conceptual replication of Llanos et al. (2020) using unfamiliar phonemic contrasts rather than lexical tones. In doing so, we hoped to clarify and extend the previous equivocal findings on taVNS and language learning. To this end, 45 native English speakers were trained on a perceptual labeling task for German front rounded vowels and fricatives. During training, 15 of the participants received stimulation paired with the vowels ("easier" phonemic contrast), 15 received stimulation paired with the fricatives ("harder" phonemic contrast),

and 15 received no stimulation (control group). The second and third objectives involved examining the potential impacts of taVNS on elements of language learning that have thus far remained unexplored. To accomplish this, participants completed a German speech production task and a motivation questionnaire pre- and post-training. We anticipated that taVNS would enhance perception and production of the unfamiliar contrasts relative to the control group, and that its effects would be greatest for vowel learning since the vowel contrast is more perceptually salient than the consonant one. We also predicted that participants in the taVNS groups would show greater increases in language learning motivation from pre to post compared to controls.

2 Materials and methods

2.1 Participants

Forty-five adults (33 females, 10 males, two preferring not to answer) were recruited from the Montreal area. This number was chosen based on the similar work of Llanos et al. (2020), who recruited 36 participants; they had 12 participants per group, and we decided to obtain 15 per group in order to try to replicate and extend their results. All participants identified as monolingual English speakers and were unfamiliar with German. Participants were aged 18–35 (mean: 23.0) with normal hearing thresholds in both ears as determined by an audiometric screening, and with no history of literacy, language, or cognitive impairments. People with medical implants, with metal braces, or who were pregnant were excluded for safety reasons. At the beginning of the experiment, participants were randomly assigned to one of three groups: taVNS-vowel ($N = 15$), taVNS-fricative ($N = 15$), and Control ($N = 15$). Participants signed an informed consent form and received monetary compensation (\$40). The duration of the entire study was ~1.5 h. The research protocol was approved by the Institutional Review Board of the Faculty of Medicine and Health Sciences of McGill University.

2.2 Tasks

2.2.1 Demographic information

Participants completed a questionnaire about demographics, language history and proficiency, and musical experience, since these factors could influence speech processing. The questionnaire was adapted from the *Language history questionnaire* (LHQ 2.0; Li et al., 2013) and the *Montreal Music History Questionnaire* (MMHQ; Coffey et al., 2011). One-way ANOVAs confirmed that the extent of second language (L2) experience and musical experience did not differ significantly across groups. A summary of these ANOVAs and of demographic information for each group can be found in [Supplementary Table 1](#).

2.2.2 Non-native speech perception training

To address our first objective, participants were trained to distinguish a German consonant contrast (palatal vs. postalveolar

TABLE 1 German minimal pairs used in the non-native speech perception training task.

Consonant contrast		Vowel contrast	
Palatal fricative (ç)	Postalveolar fricative (ʃ)	Tense high front rounded vowel (y)	Lax high front rounded vowel (ʏ)
Fichte /fɪçtə/	fischte /fɪʃtə/	Brühl /brʏʏl/	brüll /brʏʏl/
Kirche /kɪʁçə/	Kirsche /kɪʁʃə/	Düne /dyːnə/	dünne /dyːnə/
Löchern /lœçtən/	löschern /lœʃtən/	fühlen /fyːlən/	füllen /fyllən/
selig /zeltç/	seelisch /zeltʃ/	Hüte /hyːtə/	Hütte /hyːtə/
Wicht /vɪçt/	wischt /vɪʃt/	Wüste /vyːstə/	wüsste /vyːstə/

fricative; ç vs. ʃ) and a German vowel contrast (tense vs. lax high front rounded vowel; y: vs. ʏ) which are known to be perceptually challenging sounds for native English speakers (Mayr and Escudero, 2010). English speakers tend to perceive both German ç and ʃ as English ʃ (Moulton, 1962), whereas they tend to perceive German y: and ʏ as English u: and ʊ, respectively (Strange et al., 2004; Mayr and Escudero, 2010). In line with Best's Perceptual Assimilation Model, which predicts that non-native sounds will be better discriminated when they are assimilated to two different native categories than when they are assimilated to a single native category (Best, 1994), previous work has shown that native English speakers perceive the German vowel contrast more accurately than the fricative one (Honda et al., 2024). The 10 German minimal pairs used in the training task are displayed in Table 1. To construct the stimuli, four native German speakers were recorded (two males, two females). The resulting sound files were edited to leave 20 ms before and after each production, and maximum amplitudes were normalized across speakers using GoldWave version 6.15 (GoldWave Inc, 2015). Each speaker produced each minimal pair once, resulting in a total of 80 speech stimuli (4 speakers × 2 contrasts × 10 words).

The training procedure was based on that of Llanos et al. (2020), who used a forced-choice task to present stimuli in six training blocks and one generalization block. During training, half of our German stimuli ($N = 40$, from two speakers—one male, one female) were presented in six blocks, with each stimulus being presented once per block. On each trial, participants heard a stimulus and indicated which phoneme it contained by choosing between two options via mouse click (side of the screen counterbalanced across participants). The palatal vs. postalveolar fricatives were represented by the symbols “ç”/“sh” and the tense vs. lax vowels were represented by the symbols “ü:”/“ü” to facilitate learning without needing to teach participants the International Phonetic Alphabet. Visual feedback (“Correct”/“Incorrect”) was provided immediately after each trial. Feedback lasted 1,000 ms, and there were 500 ms between the end of feedback and the onset of the following stimulus. After the six training blocks, participants completed a Generalization block during which they labeled the other half of the stimuli ($N = 40$, from the other two speakers). There was no feedback or stimulation during the Generalization block. To avoid physical interference with the stimulation

electrodes placed on the left ear, audio was delivered monaurally through the right ear with an insert earphone. The Training and Generalization blocks were programmed and presented using E-Prime 3.0.

2.2.3 Electrical stimulation procedure

Transcutaneous stimulation of the vagus nerve occurred during the perception training task. Replicating the procedure used in Llanos et al. (2020), stimulation was delivered through the cymba concha and cavum concha of the outer ear, at a level below each participant's perceptual threshold (as described further below). The participant's left ear was first cleaned with an alcohol swab. Next, silicon putty was molded to the shape of the participant's ear. The molded putty had an indentation across the middle caused by the crus of the helix, demarcating the cymba concha and cavum concha on either side. Two Ag-AgCl disc electrodes were then embedded in the putty in the center of the areas corresponding to the cymba concha (cathode) and cavum concha (anode) and covered with conductive gel. Finally, the mold was pressed into place in the ear. The same experimenter performed the electrode setup for all participants to ensure maximal consistency in the procedure. Electrical stimulation was generated with a BIOPAC STMISOLA Constant Current Isolated Linear Stimulator. Consistent electrode contact was ensured by monitoring the stimulator's “Protect” light, which turns on only when contact is lost (this occurred on occasion during electrode setup and calibration, in which case the setup steps were repeated, but did not occur during training). Stimulation waveforms consisted of 14 biphasic square-wave pulses (150 μ s pulse width), delivered at a rate of 25 Hz and with an amplitude no higher than 3 mA for safety reasons. The pulse train began at the onset of the auditory stimulus and continued for 560 ms. These stimulation parameters were selected based on Llanos et al. (2020) who found significant taVNS effects on speech sound learning using the same pulse width, frequency, and amplitude specifications, and they also closely resemble the parameters used in other work (e.g., Engineer et al., 2015; McHaney et al., 2023). Pulses were generated using Matlab (Mathworks, 2017) and transmitted to the stimulator via a Measurement Computing USB-1208HS DAQ card.

Before the non-native speech perception training, each participant’s perceptual threshold for the taVNS was identified through a calibration procedure. During calibration, individual pulse trains were delivered with the same parameters described above, starting at 0.1 mA and increasing in steps of 0.2 mA until the participant indicated feeling the stimulation. Amplitude was then decreased in steps of 0.1 mA until the participant no longer felt the stimulation. Each participant’s threshold was recorded as the amplitude at which they could reliably begin to feel the stimulation across at least two repetitions of this procedure. During training, stimulation was delivered with a pulse amplitude 0.2 mA below the participant’s perceptual threshold. There were no significant differences in pulse amplitude between the two groups that received stimulation (taVNS-vowel: $M = 0.61$ mA, $SD = 0.35$ mA; taVNS-fricative: $M = 0.67$ mA, $SD = 0.45$ mA; two-sample t -test: $t_{26} = -0.41, p = 0.69$). The control group underwent the same setup and threshold determination procedures so that all participants were blind to the condition to which they were assigned.

2.2.4 Non-native speech production task

To address our second objective, participants completed a non-native speech production task before and after the speech perception training. We wanted to test how well the perceptual training transferred to participants’ general ability to produce the trained sounds accurately, and not simply their ability to imitate words they had been trained on. We therefore used a consonant-vowel-consonant (CVC) syllable production task and compared pre- and post-training productions. Each trial of the task involved a familiarization component followed by a production component. Participants were first familiarized with a non-native speech sound by hearing it presented in an isolated syllable (/çə/ and /ʃə/ for the fricatives, /y/ and /ɪ/ for the vowels). Speech sounds were produced by a third male native German speaker, different from the ones used in training. In order to reduce transfer of learning between the production task and the subsequent perceptual training task, we tried to maximize differences between the stimuli by using isolated syllables and a different voice (Bradlow et al., 1997; Baker and Trofimovich, 2006). The auditory presentation of the

non-native speech sounds was accompanied by visual presentation of letters on a screen so that participants learned to associate each sound with a simplified version of its corresponding orthography (palatal vs. postalveolar fricatives represented by “ç”/“sh” and tense vs. lax vowels represented by “ü”/“u” as described in the training task). In this way, the production task also served to familiarize participants with the phonemes that would subsequently be trained. On each trial, after being familiarized with the native exemplar of the speech sound, participants were prompted to produce the sound within a CVC syllable. There were six different CVCs to produce for each sound (6×4 sounds = 24 productions total), and CVCs were written using English orthography combined with the letters that had been associated with the German sounds (ç, sh, ü, u). Trials were blocked so that all six trials for a given sound appeared together, and block order was randomized. Table 2 displays the full list of CVCs used. Participants’ productions were recorded using a headset microphone (Logitech, Switzerland). For each participant, all 24 productions were recorded before and after the training phase. The production task was programmed and presented using E-Prime 3.0.

TABLE 3 Intrinsic Motivation Inventory (IMI), modified for the current study.

Interest/enjoyment
I enjoy foreign language exercises very much.
I think foreign language exercises are boring activities. (R)
I would describe foreign language exercises as very interesting.
Foreign language exercises do not hold my attention at all. (R)
Perceived competence
I think I am pretty good at learning foreign languages.
After practicing foreign languages for awhile, I feel pretty competent.
I am satisfied with my performance at learning foreign languages.
Foreign language exercises are an activity that I can’t do very well. (R)
Tension/pressure
I feel very tense while practicing foreign languages.
I am very relaxed in practicing foreign languages. (R)
I do not feel nervous at all while practicing foreign languages. (R)
I feel pressured while practicing foreign languages.

An R after an item indicates a reverse item.

TABLE 2 German CVCs used in the non-native speech production task.

Consonant contrast		Vowel contrast	
Palatal fricative (ç)	Postalveolar fricative (ʃ)	Tense high front rounded vowel (y)	Lax high front rounded vowel (ɪ)
kiç /kiç/	kish /kiʃ/	hü:n /hy:n/	hün /hɪn/
liç /liç/	lish /liʃ/	pü:s /py:s/	pün /pɪs/
geeç /giç/	geesh /giʃ/	fü:m /fy:m/	fün /fɪm/
veeç /viç/	veesh /viʃ/	mü:p /my:p/	mün /mɪp/
deç /deç/	desh /deʃ/	lü:m /ly:m/	lün /lɪm/
tayç /teiç/	taysh /teiʃ/	kü:t /ky:t/	kün /kɪt/

Recordings of participants' productions were subsequently presented to three native German speakers. The native speakers completed ratings of the recordings at home on Gorilla Experiment Builder (www.gorilla.sc; Anwyl-Irvine et al., 2020), using their own headphones. Productions were presented in three sessions of ~1 h each; all three sessions were completed within 1 week, with a minimum break of 1 h between sessions to avoid fatigue. Within each session, productions from one third of participants (five per group) were randomly mixed, and productions from pre- and post-training were randomly mixed. For each production, the native speakers indicated which sound they heard via mouse click [2-alternative forced choice (2AFC) between the two sounds that make up the contrast]. In addition, the three native speakers rated the quality of each production using a 7-point Likert scale (1 = poor, 7 = native-like). These quality ratings provided a more fine-grained measure of pronunciation ability. While acoustic analyses could also have been used to rate participants' productions, native speaker ratings were chosen because they provide a global accuracy measure that accounts for a variety of acoustic and articulatory dimensions which would be difficult to examine individually in isolation.

2.2.5 Language learning motivation

To address our third objective, we measured participants' motivation to learn foreign languages using a modified version of the Intrinsic Motivation Inventory (IMI; McAuley et al., 1989). Four items from three subscales (Interest/Enjoyment, Perceived Competence, Tension/Pressure) were used (following Saito, 2021), as shown in Table 3. Participants indicated how true the items were for them on a 7-point Likert scale (1 = not at all true, 7 = very true). Higher scores represent higher motivation for Interest/Enjoyment and Perceived Competence, while lower scores represent higher

motivation for Tension/Pressure. The IMI was administered pre- and post-training.

3 Analyses and results

The raw data and all code used to process and analyze it is publicly available on the OSF (https://osf.io/fdsaz/?view_only=d4f9ea6ef9804606b4972b3488981dc3). Details and output of all analyses described below can be found in the R Markdown document on the same OSF page. For all regression models reported in this experiment, the maximal random effects structure was used (Barr et al., 2013) unless the model failed to converge, in which case random effects were removed one at a time until model convergence was achieved. Note that in cases where our analyses were following previous work by Llanos et al. (2020) and McHaney et al. (2023), our maximal model is reported in the main text but an additional model with exactly the same structure as in the previous work was also fit for the sake of replication. The additional models' output can be found in the R Markdown document. All analyses were carried out in R (R Development Core Team, 2020): mixed-effects logistic models were fit using the package *lme4* (Bates et al., 2015), linear mixed-effects models were fit using the package *lmerTest* (Kuznetsova et al., 2017), and ordinal mixed-effects models were fit using the package *ordinal* (Christensen, 2022).

3.1 Effects of stimulation on perception

3.1.1 Accuracy improvement during training

3.1.1.1 Analyses

To assess the potential effects of taVNS on the learning of non-native sounds, a mixed-effects logistic regression model

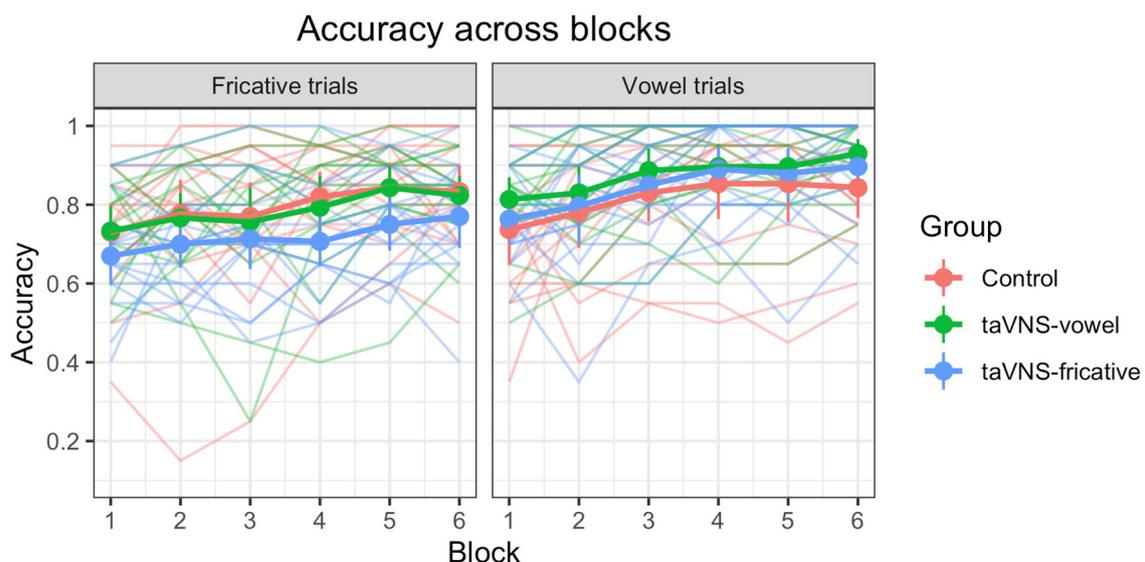


FIGURE 1

Accuracy over time on the training task (from the first to the last block), for individuals and groups. Thin translucent lines represent individual participants' data, whereas thick solid lines represent aggregate group data. Vertical lines denote 95% confidence intervals for group accuracy on each block. Note the high levels of individual variability. Overall, accuracy improved over the course of training for all groups, and there were no significant group differences.

TABLE 4 Summary of the mixed-effects logistic regression model predicting accuracy on the training task.

Fixed effects				
Coefficient	$\hat{\beta}$	SE ($\hat{\beta}$)	z	p
Intercept	1.723	0.200	8.614	<0.001
Group taVNS-vowel	0.167	0.283	0.590	0.555
Group taVNS-fricative	-0.178	0.282	-0.633	0.527
Trial number	0.619	0.121	5.107	<0.001
Contrast	-0.296	0.271	-1.094	0.274
Group taVNS-vowel:trial number	0.073	0.173	0.424	0.672
Group taVNS-fricative:trial number	-0.004	0.168	-0.026	0.979
Group taVNS-vowel:contrast	-0.582	0.384	-1.515	0.130
Group taVNS-fricative:contrast	-0.674	0.381	-1.769	0.077
Trial Number:contrast	-0.062	0.218	-0.286	0.775
Group taVNS-vowel:trial number:contrast	-0.36	0.315	-1.141	0.254
Group taVNS-fricative:trial number:contrast	-0.461	0.304	-1.515	0.130
Random effects				
Group	Term	Variance	SD	
Subject	Intercept	0.555	0.745	
	Trial number	0.079	0.280	
	Contrast	0.929	0.964	
	Trial number:contrast	0.172	0.414	

Number of observations: 10,800, groups: subject (45). *p*-values calculated using the Laplace approximation. Model equation: accuracy ~ group * trial number * contrast + (1 + trial number * contrast || subject).

was fit similarly to that found in Llanos et al. (2020). The dependent variable was trial-level responses (correct/incorrect) for each participant during the training blocks. Fixed effects consisted of group (taVNS-vowel, taVNS-fricative, and Control = reference level), trial number (1–240; centered and divided by 2 SD), and contrast (fricative = 0.5, vowel = -0.5), along with all two- and three-way interactions among those three variables. For this analysis, the maximal model that converged included by-subject random intercepts and by-subject random slopes of trial number, contrast, and the interaction between trial number and contrast, without correlations between random effects. The group-by-trial interaction revealed whether the taVNS groups showed greater improvement over the course of training compared to the control group.

Following McHaney et al. (2023), another mixed-effects logistic regression model was fit to determine whether participants' performance during the training blocks depended on the amplitude of taVNS received or the type of trial (stimulated vs. unstimulated). The dependent variable was trial-level accuracy for participants in the two stimulation groups. There were fixed effects of trial type (stimulated = 0.5, unstimulated = -0.5), trial number, amplitude (centered and divided by 2 SD), and all two- and three-way

interactions among these variables. Random effects consisted of by-subject and by-stimulus random intercepts, by-subject random slopes of trial type and trial number, and by-stimulus random slopes of trial number and amplitude, without correlations between random effects. A linear mixed-effects regression model was also fit predicting the retention of correct stimulus-response associations across blocks (as done in Llanos et al., 2020), to determine whether stimulation improved retention over time. The dependent variable for this model was the percentage of trials correctly labeled on both the current block and the previous block, starting at block 2. Fixed effects consisted of group, block (2–6; block 2 = reference level), and contrast, as well as all two- and three-way interactions among them. Random effects consisted of by-subject random intercepts and by-subject random slopes of contrast, without correlations between random effects.

3.1.1.2 Results

Figure 1 displays accuracy over the course of training, at the individual and group level. We can see that accuracy improved for all three groups and for both speech sounds over the course of training. For the model predicting trial-level accuracy during training, there was a significant effect of trial number, indicating that the control group improved their performance over time across both contrasts ($\hat{\beta} = 0.619, p < 0.001$); we can conclude that training resulted in learning even when no stimulation was administered. No other significant effects were found; see Table 4 for a full model summary. Importantly, there were no significant group-by-trial number interactions; this demonstrates that although all three groups showed improved performance over time, the stimulation groups did not show greater improvement compared to the control group.

For the model including stimulation amplitude and trial type as fixed effects, there was again only a significant effect of trial number ($\hat{\beta} = 0.684, p < 0.001$); performance on the perception task improved during training regardless of the stimulation condition, and these improvements in performance did not depend on stimulation amplitude. The lack of a significant main effect of trial type (stimulated vs. unstimulated) also indicates that stimulation failed to improve training performance. See Supplementary Table 2 for full model output. Supplementary Figure 1 shows the similar accuracy trajectory over time for stimulated and unstimulated trials. For the model predicting retention of correct stimulus-response associations, stimulation did not significantly predict retention rates. The only significant predictor was block ($p = 0.011$ for block 3 retention compared to block 2 retention, $p < 0.005$ for the other three blocks compared to block 2 retention), indicating better retention as training progressed (i.e., learning). See Supplementary Table 3 for model output and Supplementary Figure 2 for each group's retention rates across blocks. Overall, these analyses converge on the conclusion that performance on the training task improved over time but was not affected by stimulation.

3.1.2 Reaction times during training

3.1.2.1 Analyses

Given that taVNS has in some cases been shown to decrease reaction times (RTs; Pandža et al., 2020; Chen et al., 2021; Phillips et al., 2021) and increase post-error slowing (PES; Sellaro

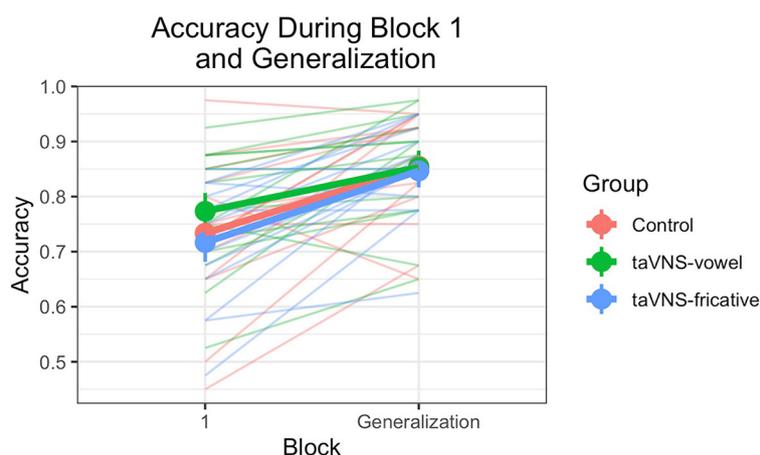


FIGURE 2

Accuracy on block 1 of training and on the generalization block, for individuals and groups. Thin translucent lines represent individual participants' data, and thick solid lines represent aggregate group data. Vertical lines denote 95% confidence intervals for group accuracy on each block. Note the high levels of individual variability. Overall, accuracy improved from block 1 to the generalization block, and there were no significant group differences.

et al., 2015), analyses were run to determine the potential effects of stimulation on RTs and PES during training. For the RT analysis, trials on which participants responded incorrectly were removed (19% of trials) along with trials where RTs were < 200 ms or > 2.5 SD above the participant's mean (a further 2% of trials), following Giannakopoulou et al. (2017). The distribution of raw RT values was positively skewed, so RTs were log transformed. A mixed-effects linear regression model was fit with the resulting cleaned and transformed RTs as the dependent variable. Fixed effects consisted of group, trial number, contrast, and all two- and three-way interactions between them. Random effects consisted of by-stimulus and by-subject random intercepts, as well as by-subject random slopes of trial number, contrast, and the interaction between trial number and contrast, without correlations between random effects. To calculate PES values, the RT on a trial following an error was subtracted from the RT on the trial preceding that error. A linear mixed-effects regression model was then fit predicting PES values, with group as a fixed effect and by-stimulus random intercepts as the maximal random effects structure promoting convergence and non-singularity.

3.1.2.2 Results

For the RT model, the only significant predictor was trial number ($\hat{\beta} = -233.656$, $p < 0.001$), indicating that participants in the control group became faster at responding to both contrasts as the task progressed. The lack of group-by-trial number interaction indicates that stimulation did not affect this decrease in reaction times. See Supplementary Table 4 for model output and Supplementary Figure 3 for each group's RTs plotted against trial number. For the PES model, no significant predictors were found; stimulation did not increase post-error slowing. See Supplementary Table 5 for model output and Supplementary Figure 4 for PES values per group.

3.1.3 Accuracy during generalization block

3.1.3.1 Analyses

As in Llanos et al. (2020), a mixed-effects logistic regression model was fit with the dependent variable being trial-level accuracy in the generalization block and in block 1 of training. Fixed effects were group, block (generalization, block 1 = reference level), and contrast, along with all two- and three-way interactions between them. Random effects consisted of by-stimulus and by-subject random intercepts, along with by-subject random slopes of block, contrast, and the interaction between block and contrast. The group-by-block interaction enabled us to determine whether stimulation groups showed better generalization of their learning after accounting for baseline performance during block 1 of training.

3.1.3.2 Results

Accuracy during block 1 and the generalization block is displayed in Figure 2. Table 5 shows the output of the model predicting accuracy during block 1 and the generalization block. There were no significant effects of group in block 1, demonstrating that groups did not differ in baseline performance across contrasts at the beginning of training. Block was the only significant predictor ($\hat{\beta} = 1.022$, $p < 0.001$), indicating that performance was better during the generalization block than during block 1 of training—in other words, participants successfully learned and generalized their learning to new voices, regardless of group.

3.2 Effects of stimulation on production

3.2.1 Analyses

Interrater reliability measures were obtained for the native German speakers' ratings of the production data. Fleiss' Kappa was calculated for the 2AFC ratings and the intraclass correlation coefficient (ICC; two-way random-effects model) was calculated for the Likert ratings (Gisev et al., 2013; Koo and Li, 2016).

These calculations revealed acceptable reliability for both rating types (Fleiss' Kappa = 0.384; ICC = 0.594) based on established guidelines ("fair" to "substantial" according to Landis and Koch, 1977).

For the 2AFC ratings, a mixed-effects logistic regression model was fit with trial-level accuracy as the dependent variable. Fixed effects were group, time (post, pre = reference level), contrast, and all two- and three-way interactions among them, as well as rater (first rater = reference level). Random effects consisted of by-subject random intercepts and by-subject random slopes of time, contrast, and the interaction between time and contrast, without correlations between random effects. The group-by-time interaction revealed whether stimulation groups showed greater increases in production accuracy from pre to post compared to the control group.

For the Likert ratings, an ordinal mixed-effects regression model was fit. Fixed effects were the same as for the model predicting 2AFC ratings. Random effects consisted of by-subject random intercepts and by-subject random slopes of time, contrast, and the interaction between time and contrast. The group-by-time interaction revealed whether native speakers' ratings of the productions increased more from pre to post for the stimulation groups compared to the control group.

3.2.2 Results

Figure 3 shows the differences in ratings of productions from pre- to post-training, for the 2AFC ratings (Figure 3A) and the Likert ratings (Figure 3B). For the model predicting 2AFC ratings, rater was a significant predictor, with raters 2 and 3 tending to rate pre-training productions from participants in the control group more accurately compared to rater 1 ($\beta = 0.162, p = 0.014$ for rater 2; $\beta = 0.215, p = 0.001$ for rater 3). Table 6 displays the full output of the model. For the model predicting Likert ratings, rater was again a significant predictor, with rater 3 giving higher ratings than rater 1 ($\beta = 0.785, p < 0.001$) for pre-training productions from control group participants. Contrast was also a significant predictor, with rater 1 giving higher ratings to fricatives than to vowels ($\beta = 0.612, p < 0.001$). This effect was driven by high ratings for the fricative *ʃ* since it also occurs in English. Finally, time was also a significant predictor ($\beta = 0.230, p = 0.005$) and did not interact with other predictors, revealing that ratings of production quality increased from pre- to post-training regardless of group. Full model output can be found in Table 7. Across both models, the lack of group effects or of group-by-time interactions suggests that the administration of taVNS during perceptual training did not specifically improve the subsequent production of the trained sounds.

3.3 Effects of stimulation on motivation

3.3.1 Analyses

To determine the potential effects of taVNS on language learning motivation, ordinal mixed-effects models were fit as in Saito (2021). Fixed effects consisted of group, time (post, pre = reference level), and the interaction between the two, while random

TABLE 5 Summary of the mixed-effects logistic regression model predicting accuracy on block 1 and the generalization block.

Fixed effects					
Coefficient	$\hat{\beta}$	SE ($\hat{\beta}$)	z	p	
Intercept	1.213	0.224	5.418	<0.001	
Group taVNS-vowel	0.244	0.259	0.941	0.347	
Group taVNS-fricative	-0.115	0.256	-0.449	0.654	
Block generalization	1.022	0.276	3.706	<0.001	
Contrast	-0.060	0.380	-0.157	0.876	
Group taVNS-vowel:block generalization	-0.257	0.283	-0.908	0.364	
Group taVNS-fricative:block generalization	-0.003	0.278	-0.009	0.993	
Group taVNS-vowel:contrast	-0.422	0.396	-1.067	0.286	
Group taVNS-fricative:contrast	-0.472	0.388	-1.218	0.223	
Block generalization:contrast	0.202	0.524	0.385	0.700	
Group taVNS-vowel:block generalization:contrast	-0.005	0.513	-0.010	0.992	
Group taVNS-fricative:block generalization:contrast	0.056	0.502	0.111	0.912	
Random effects					
Group	Term	Variance	SD	Correlation	
Stimulus	Intercept	0.661	0.813		
Subject	Intercept	0.341	0.584		
	Block generalization	0.196	0.443	-0.15	
	Contrast	0.528	0.727	-0.21	0.23
	Block generalization:contrast	0.368	0.607	0.32	-0.24

Number of observations: 3,600, groups: stimulus (80), subject (45), p-values calculated using the Laplace approximation. Model equation: accuracy ~ group * block * contrast + (1 + block * contrast | subject) + (1 | stimulus).

effects consisted of by-item random intercepts, by-subject random intercepts, and by-subject random slopes of time. The group-by-time interaction revealed whether stimulation groups showed greater motivation increases from pre to post compared to the control group. A first model was fit predicting responses to all items of the IMI. Three follow-up models were then fit predicting responses to each subscale: one model had an identical structure to the overall model, while the other two had only by-subject random intercepts to avoid singular fits. These models allowed us to investigate whether different aspects of motivation might be differentially affected by stimulation. Given that this follow-up analysis involved fitting more than one model to the same dataset,

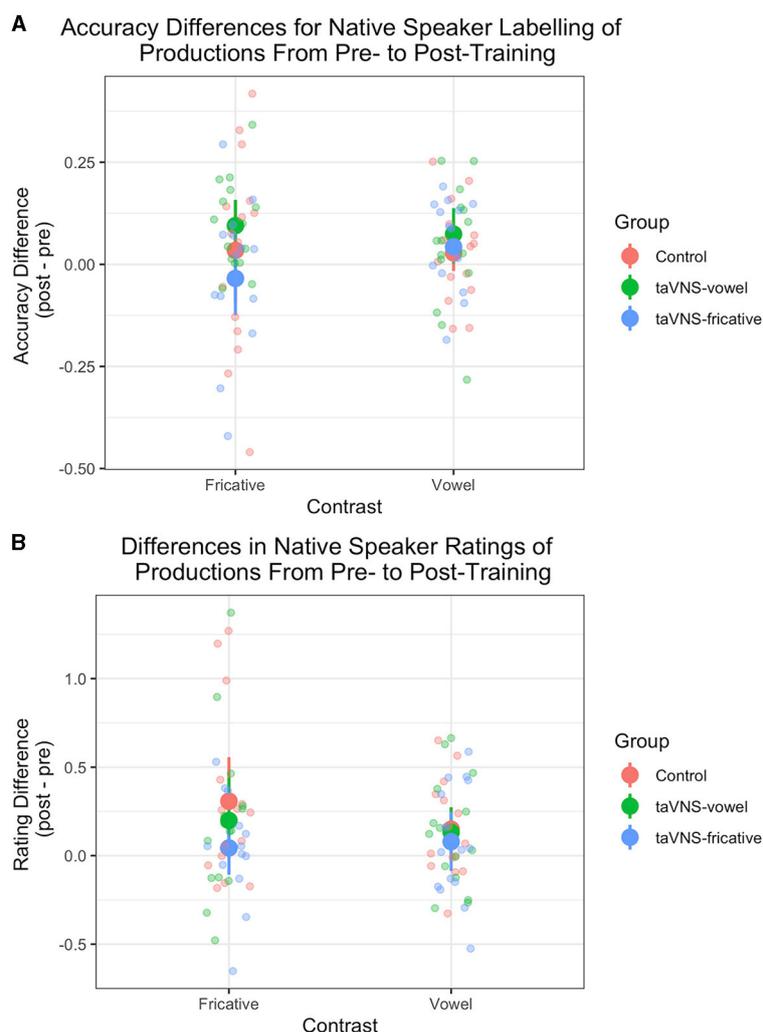


FIGURE 3

Native German speakers' ratings of participants' productions pre- and post-training. (A) Accuracy on a 2-alternative forced choice task. (B) Ratings on a 7-point Likert scale (1 = poor, 7 = native-like). Thin translucent lines represent individual participants' data, whereas thick solid lines represent aggregate group data. Vertical lines denote 95% confidence intervals. There was a significant increase in Likert ratings but not in accuracy judgments of productions from pre- to post-training, and there were no significant differences between groups.

Bonferroni corrections were performed on the resulting p -values by dividing alpha by the number of comparisons being made ($0.05/3 = 0.017$).

3.3.2 Results

Individual- and group-level motivation scores on each subscale pre- and post-training are displayed in Figure 4. As was the case for the perception and production measures, considerable individual variability in motivation scores was found. There were no significant predictors for the model predicting responses across all subscales (Supplementary Table 6). For the model predicting responses on the Tension/Pressure subscale, there was a group-by-time interaction such that the taVNS-vowel group showed greater increases in scores (indicating reduced feelings of tension and pressure) compared to the control group ($\hat{\beta} = 1.042$, $p = 0.029$). Although this interaction was not significant after Bonferroni

correction ($p < 0.017$), *post-hoc* pairwise comparisons between each group's scores pre- and post-training revealed that the taVNS-vowel group did show a significant difference from pre to post ($\hat{\beta} = -0.846$, $p < 0.001$), which was not the case for the other groups. Stimulation during easier-to-learn sounds may therefore have decreased participants' feelings of tension and pressure associated with language learning. Output of the model and of the *post-hoc* comparisons can be found in Table 8.

4 Discussion

The goal of the present study was to investigate whether the administration of taVNS during non-native speech perception training could enhance (1) participants' perception of the trained sounds, (2) their subsequent production of those same sounds, and (3) their language learning motivation. Native English speakers underwent training to perceive unfamiliar German

TABLE 6 Summary of the mixed-effects logistic regression model predicting accuracy of native speakers' ratings of participants' productions.

Fixed effects				
Coefficient	$\hat{\beta}$	SE($\hat{\beta}$)	z	p
Intercept	0.432	0.117	3.697	<0.001
Group taVNS-vowel	-0.027	0.157	-0.174	0.862
Group taVNS-fricative	0.194	0.157	1.234	0.217
Time post	0.19	0.124	1.532	0.125
Contrast	0.318	0.229	1.391	0.164
Rater 2	0.162	0.066	2.454	0.014
Rater 3	0.215	0.066	3.237	0.001
Group taVNS-vowel:time post	0.282	0.177	1.589	0.112
Group taVNS-fricative:time post	-0.153	0.176	-0.871	0.384
Group taVNS-vowel:contrast	0.038	0.324	0.116	0.908
Group taVNS-fricative:contrast	0.016	0.325	0.050	0.960
Time post:contrast	0.125	0.279	0.446	0.655
Group taVNS-vowel:time post:contrast	0.165	0.398	0.415	0.678
Group taVNS-fricative:time post:contrast	-0.487	0.396	-1.232	0.218
Random effects				
Group	Term	Variance	SD	
Subject	Intercept	0.120	0.347	
	Time	0.097	0.311	
	Contrast	0.532	0.729	
	Time:contrast	0.632	0.795	

Number of observations: 6,450, groups: subject (45). *p*-values calculated using the Laplace approximation. Model equation: accuracy ~ group * time * contrast + rater + (1 + time * contrast || subject).

vowels and fricatives, and were randomly assigned to one of three groups: taVNS-vowel (stimulation during easier-to-perceive sounds), taVNS-fricative (stimulation during harder-to-perceive sounds), or Control (no stimulation). Participants completed a German speech production task and a language learning motivation questionnaire before and after training. Contrary to our expectations, we did not find clear benefits of taVNS for perception, production, or motivation. However, it is possible that taVNS during training may have specifically improved language learning motivation related to pressure/tension. In particular, the taVNS-vowel group showed reduced feelings of pressure from pre- to post-training, indicating that stimulation during easier-to-perceive sounds may alleviate certain negative feelings associated with learning a new language in adulthood. The fact that taVNS was found to affect feelings of pressure but not the other motivation subscales measured (enjoyment and competence) is in line with previous work showing that taVNS prior to language learning tasks decreased negative affect but did not increase positive affect (Calloway et al., 2020). Nevertheless, as elaborated upon below, the

TABLE 7 Summary of the ordinal mixed-effects regression model predicting native speakers' Likert ratings of participants' productions.

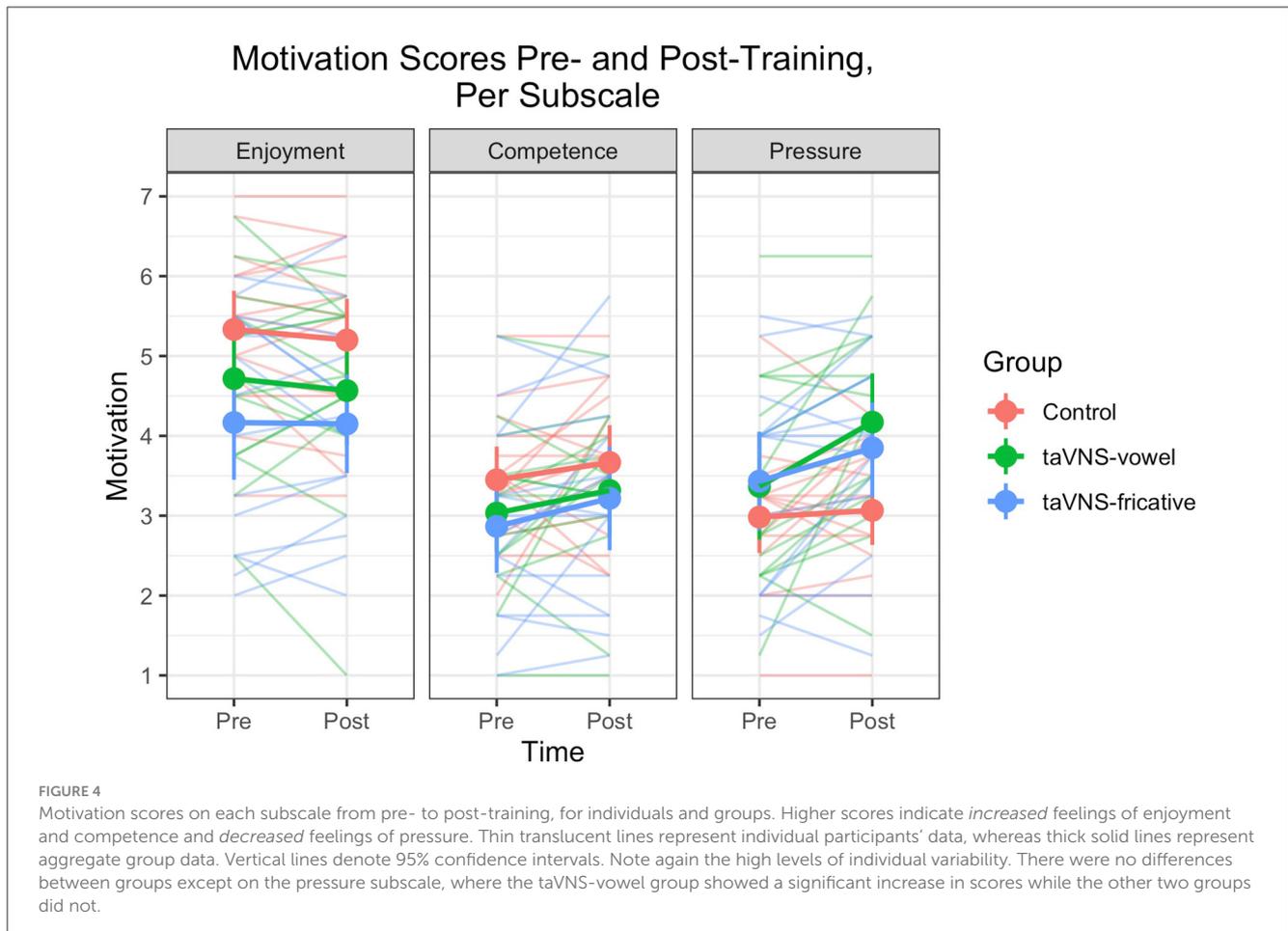
Fixed effects					
Coefficient	$\hat{\beta}$	SE($\hat{\beta}$)	z	p	
Intercept 1 2	-1.279	0.113	-11.284	<0.001	
Intercept 2 3	-0.301	0.112	-2.680	0.007	
Intercept 3 4	0.418	0.113	3.712	<0.001	
Intercept 4 5	1.150	0.113	10.142	<0.001	
Intercept 5 6	1.922	0.115	16.756	<0.001	
Intercept 6 7	3.187	0.121	26.362	<0.001	
Group taVNS-vowel	0.147	0.150	0.979	0.327	
Group taVNS-fricative	0.146	0.150	0.973	0.330	
Time post	0.230	0.081	2.824	0.005	
Contrast	0.785	0.168	4.667	<0.001	
Rater 2	-0.012	0.056	-0.208	0.835	
Rater 3	0.612	0.054	11.360	<0.001	
Group taVNS-vowel:time post	-0.056	0.114	-0.492	0.622	
Group taVNS-fricative:time post	-0.176	0.115	-1.532	0.125	
Group taVNS-vowel:contrast	-0.064	0.237	-0.270	0.787	
Group taVNS-fricative:contrast	-0.228	0.237	-0.966	0.334	
Time post:contrast	0.198	0.165	1.203	0.229	
Group taVNS-vowel:time post:contrast	-0.102	0.232	-0.438	0.661	
Group taVNS-fricative:time post:contrast	-0.224	0.233	-0.964	0.335	
Random effects					
Group	Term	Variance	SD	Correlation	
Subject	Intercept	0.122	0.349		
	Time	0.008	0.087	0.922	
	Contrast	0.232	0.482	-0.228	-0.588
	Time:contrast	0.040	0.201	0.888	0.640

Number of observations: 6,450, groups: subject (45). *p*-values calculated using the Laplace approximation. Model equation: response ~ group * time * contrast + rater + (1 + time * contrast | subject).

potential benefits of taVNS for language learning and motivation remain largely uncertain; further research is called for to clarify and extend these findings.

4.1 Lack of evidence for taVNS-related improvement in non-native phonetic perception

While we had hypothesized that taVNS might improve the perception of non-native phonemic contrasts, there are several



potential reasons why this did not end up being the case. Given that music experience has been shown to predict the success of non-native sound learning (e.g., Slevc and Miyake, 2006; Perfors and Ong, 2012), one potential concern is that differences in our groups' musical backgrounds could have influenced our results. While an ANOVA revealed no significant differences between groups as mentioned earlier, the taVNS-fricative group had the fewest years of music training and the control group had the most (see Supplementary Table 1). To verify whether these numeric differences could be playing a role in the efficacy of stimulation, a mixed-effects logistic regression model was fit. The model had the same structure as the aforementioned model predicting accuracy during training, but included years of music training (centered and divided by 2 SD) as an additional fixed effect. As with the model that did not include music training, trial number was the strongest predictor, indicating learning over time ($\hat{\beta} = 5.195$, $p < 0.001$); and music training was a significant predictor ($\hat{\beta} = 2.376$, $p = 0.018$), pointing to the benefits of music experience for speech sound learning (see R Markdown document on the OSF page for full model output). However, no significant main effects of group or group-by-trial interactions were found; even after accounting for music training, taVNS did not improve overall accuracy or accelerate learning across trials.

An additional concern is that participants generally showed quite accurate performance on the training task, even from the beginning (see Figure 1). It is possible that stimulation-related effects were not observed because there was limited room for improvement in performance. To partly alleviate this concern, an additional mixed-effects logistic regression model was fit with the same structure as the main model predicting accuracy during training, but with block 1 accuracy (centered and divided by 2 SD, and excluding accuracy on block 1 trials) as an additional fixed effect. As with the other training models, trial number was the strongest predictor ($\hat{\beta} = 3.939$, $p < 0.001$), and there were no significant main effects of group or group-by-trial interactions (see R Markdown document). Thus, although many participants had limited room for improvement during training, it does not appear that any potential effects of stimulation depended on initial accuracy.

Having addressed these concerns through additional analyses, we may turn to other possible explanations for our null findings. Previous work with taVNS and speech sound learning has focussed only on tonal contrasts (Llanos et al., 2020; Pandža et al., 2020; Phillips et al., 2021; McHaney et al., 2023). It is possible that taVNS does not benefit the learning of phonemic contrasts in the same way that it may benefit the learning of tonal contrasts. Evidence suggests that during the processing of novel phonemic

TABLE 8 Summary of the ordinal mixed-effects regression model and *post-hoc* pairwise comparisons predicting participants' motivation ratings on the pressure/tension subscale.

Fixed effects				
Coefficient	$\hat{\beta}$	SE($\hat{\beta}$)	z	p
Intercept 1 2	-2.074	0.496	-4.185	<0.001
Intercept 2 3	-0.359	0.473	-0.759	0.448
Intercept 3 4	1.003	0.473	2.120	0.034
Intercept 4 5	2.003	0.481	4.165	<0.001
Intercept 5 6	3.250	0.503	6.464	<0.001
Intercept 6 7	4.781	0.564	8.473	<0.001
Time post	0.069	0.337	0.205	0.837
Group taVNS-vowel	0.578	0.657	0.879	0.379
Group taVNS-fricative	0.545	0.656	0.831	0.406
Time post:group taVNS-vowel	1.042	0.478	2.181	0.029
Time post:group taVNS-fricative	0.560	0.477	1.174	0.240
Random effects				
Group	Term	Variance	SD	
Subject	Intercept	2.323	1.524	
Post-hoc pairwise comparisons				
Contrast	$\hat{\beta}$	SE($\hat{\beta}$)	z ratio	p
Pre—post (control group)	-0.048	0.232	-0.205	0.837
Pre—post (taVNS-vowel group)	-0.846	0.253	-3.343	<0.001
Pre—post (taVNS-fricative group)	-0.472	0.252	-1.875	0.061

Number of observations: 358, groups: subject (45). *p*-values calculated using the Laplace approximation. Model equation: rating \sim time * group + (1 | subject).

contrasts, learners show neural activity in the same regions of the left hemisphere that are activated by native phonemic contrasts (Golestani and Zatorre, 2004). In contrast, during the processing of novel tonal contrasts, learners show neural activity in regions of the right hemisphere that are commonly associated with non-linguistic pitch processing (Hsieh et al., 2001). Perhaps our results differ from those of previous taVNS and speech sound learning studies due to differences in how novel tonal vs. phonemic contrasts are processed. Furthermore, it is likely that native English speakers show different perceptual assimilation patterns for the phonemic contrasts used in the current study compared to the tonal contrasts used in previous work (Strange et al., 2004; So and Best, 2014); differences in assimilation could affect the discriminability of the contrasts (Best and Tyler, 2007), leading to differing effects of taVNS on the acquisition of tonal vs. phonemic contrasts. Relatedly, in the current work it was only feasible to test participants with one language background on their acquisition of one subset of non-native phonemic contrasts. Future research could test participants with a variety of backgrounds and demographic characteristics on

their acquisition of different speech sounds in order to determine whether any potential taVNS effects may be moderated by factors such as participant language experience, age, or the relationship between the native and non-native languages. Another interesting avenue for future work would be to examine how taVNS affects measures of linguistic and non-linguistic memory and attention, both in the auditory and visual domains (e.g., pairing taVNS with reading or sign language training as well as with other non-linguistic cognitive tasks), in order to disentangle the specificity of stimulation-related effects and determine the optimal situations for taVNS use. It will be important for researchers to carefully design the methodology of future experiments with a view to detecting and differentiating the particular mechanism(s) of taVNS-related improvement that are expected to be at play (e.g., enhancements to auditory plasticity vs. arousal vs. memory).

It should also be noted that although the present study was conceived as a conceptual replication of previous work on tonal contrast learning (Llanos et al., 2020), some differences in study design may account in part for discrepancies between our findings and those of prior studies. For instance, in Llanos et al. (2020), stimulation began 300 ms prior to stimulus onset, and stimuli consisted of single syllables in which the trained speech sound reliably occurred in the same position. In the present work, mono- and disyllabic stimuli were used in which the trained speech sound could occur on either the first or the second syllable; this variability in the position of the trained sound may have reduced the salience of our non-native contrasts. Differences in stimulation timing and in stimulus salience may therefore in part explain some of the disparities between our findings and those of Llanos et al. (2020). Nonetheless, participants' overall high accuracy on our training task (as mentioned above) suggests that the non-native contrasts were salient enough to be effectively acquired. Beyond this, our training task provided participants with two response options, whereas Llanos et al. (2020) provided participants with four response options. As such, participants were more likely to respond correctly by chance in our study, which may have reduced our ability to detect taVNS-related improvements in perception because accuracy scores during training were less variable. Note, however, that Figures 1, 2 show widespread variability in performance all the same. Finally, in the current work and in Llanos et al. (2020), a single training session was administered; in contrast, in Pandža et al. (2020) and Phillips et al. (2021), two training sessions were administered on separate days. Perhaps taVNS is more effective when paired with speech sound training that spans more than a single session and that includes an opportunity for memory consolidation in between. This being said, given that even the 2-day training paradigms used in previous work bear limited resemblance to true language learning which occurs over long timescales, longitudinal work is needed to assess the effects of taVNS on more naturalistic learning.

It is plausible that taVNS does hold potential for improving speech sound learning, but that the optimal stimulation parameters for language learning have not yet been identified. Administration of taVNS entails the selection of various parameters including stimulation amplitude, pulse width, frequency, duration, and timing relative to stimulus presentation. Other work with taVNS has employed a wide range of parameters. For example, stimulation

amplitude has sometimes been much higher than in the present study (>4 mA; Jacobs et al., 2015; Liu et al., 2018; Kaan et al., 2021); frequency has ranged from 5 Hz (Thakkar et al., 2020) to 300 Hz (Phillips et al., 2021); stimulation has sometimes been delivered before (Calloway et al., 2020), during (Llanos et al., 2020; McHaney et al., 2023), or after (Clark et al., 1999) a learning task; stimulation has been administered in short targeted bursts (<1 s duration; De Ridder et al., 2014; Engineer et al., 2015; Llanos et al., 2020; McHaney et al., 2023) as well as continuously over an extended time period (>10 min duration; Ventura-Bort et al., 2018; Calloway et al., 2020; Kaan et al., 2021); and stimulation electrodes have been placed on the cymba concha and cavum concha (Llanos et al., 2020) or on the outer ear canal (Phillips et al., 2021). While we selected parameters following previous work on taVNS and language learning (Llanos et al., 2020), the field remains nascent and more research is needed to identify the most appropriate taVNS parameters for different desired outcomes. Future work could systematically compare the effects of different taVNS parameters on language learning in order to clarify whether stimulation has benefits and whether certain parameters may be more effective than others.

Another possibility is that taVNS may not be a reliable means of improving the learning of new speech sound contrasts, even tonal ones. The studies that have examined taVNS and tonal contrast learning to date have shown mixed outcomes, as reviewed in the introduction. The present study aimed to conceptually replicate the work of Llanos et al. (2020), who reported faster rates of Mandarin tone learning for participants receiving taVNS compared to controls. However, the same research group recently published a partial replication of their own work in which they did not find that tone learning rates differed significantly by experimental group (McHaney et al., 2023). Their additional exploratory analyses revealed only modest effects whereby taVNS at lower amplitudes initially increased learning rates during the training task, but without increasing overall accuracy on the task (McHaney et al., 2023). The other existing work on taVNS and tone learning comes from a research group that has administered taVNS under two different conditions (before vs. during Mandarin lexical tone training) and has tested a variety of outcome measures including accuracy and reaction times on word learning, lexical recognition, and recall tasks (Pandža et al., 2020; Phillips et al., 2021). The researchers only found stimulation-related benefits for a few of the many possible combinations of conditions and outcome measures. For example, Pandža et al. (2020) found that taVNS during (but not before) training was significantly associated with greater accuracy (but not decreased reaction times) on certain trials of the recognition task (but not the recall task). In a similar vein, Phillips et al. (2021) found that taVNS before (but not during) training was significantly associated with decreased reaction times (but not increased accuracy) on the recognition task (but not the word learning task). If taVNS were an effective method for improving tone learning, its effects might be expected to emerge in a more robust and uniform way across conditions and studies. These results also suggest that future work should carefully consider and compare different experimental conditions and outcome measures, since effects may depend on the nature of the stimulation and the tasks being administered.

Although we did not find specific benefits of taVNS for non-native speech sound perception, it is worth noting that our analyses converged on the conclusion that participants did in fact improve their perception of the non-native sounds over the course of training, regardless of group. While learning to perceive unfamiliar phonemic contrasts in adulthood is often difficult, it is clear that supervised training paradigms such as this one can facilitate the learning process, as has also been found in previous work (e.g., Bradlow et al., 1997; Iverson et al., 2005; Giannakopoulou et al., 2017; Reetzke et al., 2018). Regardless of the effectiveness of taVNS, it will be fruitful for future research to continue to investigate the optimal paradigms for training non-native perception in adults.

4.2 Lack of evidence for taVNS-related improvement in non-native phonetic production

As with non-native perception, no group effects were observed for our non-native production measures; taVNS did not specifically improve production from pre- to post-training. This outcome is perhaps not surprising given that no other studies to date have yet investigated taVNS and non-native production, and that the previous findings about taVNS and non-native perception have been mixed. In the present study, as in prior work, stimulation was delivered during perceptual training. Accordingly, any stimulation-related effects would be anticipated to emerge most notably for perceptual outcome measures; considering that no such effects on perception were found, it is unsurprising that no effects on production were found either.

As discussed above, the perceptual training task resulted in overall improvement in non-native perception regardless of experimental group. On the other hand, improvement in non-native production was less clear. Native speaker ratings of participants' productions did not improve pre- to post-training for the ratings involving a forced choice between the two sounds making up a contrast. For the more fine-grained measure where productions were rated on a seven-point scale from "poor" to "native-like", statistically significant improvement was found, though the size of the effect was not large (see Figure 3B). Our findings are in line with Sakai and Moorman's (2018) recent meta-analysis of the effects of perceptual training on non-native production. The authors found that perceptual training resulted in medium-sized effects on perception outcomes and in small effects on production outcomes (Sakai and Moorman, 2018). As such, non-native perception and production are understood to be linked, but perceptual training does not necessarily lead to reliable or significant improvements in production. It should also be noted that our production task differed in format from the perception training, which may account in part for the lack of strong improvement in production post-training. The training task involved listening to non-native words (Table 1), whereas the production task involved hearing and seeing isolated non-native phonemes as exemplars and then producing CVCs (Table 2). These differences arose because the production task was designed so that

participants would not directly repeat or imitate the exemplar and so that the stimuli would be feasible to produce for inexperienced learners. In future studies with participants who have greater non-native language experience, perception and production tasks could be made more similar in order to specifically examine the effects of perception training on production. For instance, the same words could be used as stimuli during the perceptual training task and the production task (e.g., Brosseau-Lapr e et al., 2013). Since our task did not measure spontaneous speech production, additional work will also be needed to determine the relationship between our training paradigm and more naturalistic production measures.

4.3 Evidence for taVNS-related improvement in language learning motivation

When looking across all items of our motivation questionnaire, we did not find effects of taVNS on language learning motivation. However, when focussing on the items belonging to the tension/pressure subscale of the questionnaire, we did find an effect: from pre- to post-training, the taVNS-vowel group showed a significant decrease in feelings of tension and pressure associated with language learning, which was not the case for the other two groups. Recall that the taVNS-vowel group received stimulation during easier-to-perceive (vowel) sounds. Llanos et al. (2020) also found taVNS effects specifically for the group stimulated during easier-to-perceive non-native sounds—this group showed enhanced learning over the course of non-native perception training. The authors argued that this finding emerged because taVNS increases arousal, and such modulation of arousal can specifically enhance memory consolidation for more perceptually salient stimuli (Llanos et al., 2020). While we did not find enhanced learning for the taVNS-vowel group, the fact that the group's language learning motivation increased could similarly relate to the perceptual saliency of the stimuli. Perhaps participants naturally felt more capable and relaxed when responding to the vowel trials on the training task because the vowel contrast was more perceptually salient, and so the administration of taVNS during those trials served as a reinforcement signal that modulated neural activity related to affect and reward, in turn leading to decreased feelings of tension post-training.

Prior work supports a role for taVNS in decreasing feelings of tension associated with language learning. There is preliminary evidence that taVNS may improve fear extinction after a fear conditioning task (Burger et al., 2017) and reduce spontaneous negative thoughts after a worry induction task (Burger et al., 2019). The technique may therefore have the potential to lessen participants' overall feelings of fear and of worry. taVNS has additionally been found to increase participants' confidence in their ability to perform a task successfully (Villani et al., 2019). In the particular context of language learning, administration of taVNS prior to a second language learning task has in some cases been demonstrated to reduce negative affect and anxiety (Calloway et al., 2020). All of these findings point to taVNS as a possible means of reducing the stress and tension that can be felt by adults during the language learning process.

At the neural level, these positive effects make sense given that taVNS has been shown to modulate the activity of various brain regions and networks involved in affect and motivation, including the locus coeruleus, raphe nucleus, and limbic system (Frangos et al., 2015; Yakunina et al., 2017; Badran et al., 2018). In clinical contexts, VNS is known to promote the release of serotonin from the raphe nucleus, leading to improved mood (Austelle et al., 2022). The reductions in feelings of pressure and tension observed in our taVNS-vowel group may be attributable in part to such changes in neural activity and in neurotransmitter release. Note, however, that this explanation remains speculative; it was beyond the scope of the current study to measure neural activity patterns or neurotransmitter levels. Future studies could consider including such additional measures to untangle the potential mechanisms whereby taVNS increases language learning motivation. It is also not entirely clear why taVNS would selectively reduce feelings of pressure/tension without affecting the other subscales of motivation measured here (namely, interest/enjoyment and perceived competence). More research is needed to clarify the generalizability of our findings.

In conjunction with language aptitude, motivation is known to be an important factor in predicting language learning outcomes (e.g., Gardner, 2000; D rnyei, 2001; Saito et al., 2018). Individuals who are motivated—for example, who are willing to expend effort in learning a language, who want to achieve a high level of competence in the language, and who have favorable attitudes toward the learning situation—tend to have greater non-native language achievement (Masgoret and Gardner, 2003). A meta-analysis found that, across different ages and learning environments, the correlation between motivation and second language achievement ranges from around 0.29–0.39 depending on the particular measure of achievement (Masgoret and Gardner, 2003). This effect size is considered small to medium based on Plonsky and Oswald's (2014) conventions for second language research, or medium to large based on Gignac and Szodorai's (2016) conventions for individual differences research. As such, the role of motivation in language acquisition is non-negligible, and if taVNS truly does impact motivation then this could have important repercussions for adult learners who are struggling to acquire a new language. Future work with taVNS could examine the construct of language learning motivation in greater detail, employing more extensive measures of motivation associated with conceptual frameworks such as the L2 motivational self system (D rnyei, 2009) and the socio-educational model of second language acquisition (Gardner, 2000).

5 Conclusion

In sum, we examined the potential effects of taVNS on the perception and production of non-native phonemic contrasts and on language learning motivation. taVNS had previously shown positive (but inconsistent) effects on the perception of non-native tonal contrasts (Llanos et al., 2020; Pand a et al., 2020; Phillips et al., 2021; McHaney et al., 2023), and so we sought to determine whether such effects might extend to the perception of phonemic (non-tonal) contrasts. To our knowledge, this is also the first time that taVNS has been investigated in relation to non-native phonetic production or language learning motivation. Overall, no clear

effects of taVNS on non-native perception or production emerged. Our results hint at a potential benefit of taVNS for language learning motivation—in particular, stimulation during the learning of easier-to-perceive sounds may decrease feelings of tension and pressure associated with language learning. Nevertheless, taVNS did not increase overall motivation across the three subscales of our motivation questionnaire or across our two stimulated groups, so its efficacy is not clear. On the whole, while taVNS is a promising technique with a multitude of potential applications, from treatment of epilepsy (Liu et al., 2018) to relief of tinnitus symptoms (Shim et al., 2015), its usefulness in the context of language learning remains to be determined. Research with taVNS is still just beginning to emerge, and the stimulation parameters and outcome measures used in previous work have been heterogeneous; going forward, it will be important to systematically compare a variety of stimulation conditions and language acquisition outcomes in an effort to more conclusively determine any possible uses of taVNS in language learning contexts. Improving both language acquisition and motivation is an especially important endeavor given the plurality of adults who are now learning new languages in our diverse and globalized world.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by Institutional Review Board of the Faculty of Medicine and Health Sciences of McGill University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

CH: Conceptualization, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft. NB: Conceptualization, Methodology, Supervision, Writing – review & editing. MC: Conceptualization, Methodology, Supervision,

Writing – review & editing. SB: Conceptualization, Methodology, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by a grant awarded by the Natural Sciences and Engineering Research Council of Canada (NSERC) to SB, a grant awarded by the Social Sciences and Humanities Research Council (SSHRC) to MC, and an NSERC Postgraduate Scholarship-Doctoral (PGS D) grant along with a Mitacs Accelerate Award awarded to CH.

Conflict of interest

NB is the co-founder of Revai working on developing taVNS for cognitive enhancement. While she assisted with designing the study as outlined under the Author contributions section, she was not involved in analyzing or visualizing the data.

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

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/flang.2024.1403080/full#supplementary-material>

References

- Abrahamsson, N., and Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: listener perception versus linguistic scrutiny. *Lang. Learn.* 59, 249–306. doi: 10.1111/j.1467-9922.2009.00507.x
- Anwyl-Irvine, A. L., Massonni, J., Flitton, A., Kirkham, N., and Evershed, J. K. (2020). Gorilla in our midst: an online behavioral experiment builder. *Behav. Res. Methods* 52, 388–407. doi: 10.3758/s13428-019-01237-x
- Austelle, C. W., O'Leary, G. H., Thompson, S., Gruber, E., Kahn, A., Manett, A. J., et al. (2022). A comprehensive review of vagus nerve stimulation for depression. *Neuromodulation* 25, 309–315. doi: 10.1111/ner.13528
- Badran, B. W., Dowdle, L. T., Mithoefer, O. J., LaBate, N. T., Coatsworth, J., Brown, J. C., et al. (2018). Neurophysiologic effects of transcutaneous auricular vagus nerve stimulation (taVNS) via electrical stimulation of the tragus: a concurrent taVNS/fMRI study and review. *Brain Stimul.* 11, 492–500. doi: 10.1016/j.brs.2017.12.009

- Baker, W., and Trofimovich, P. (2006). Perceptual paths to accurate production of L2 vowels: the role of individual differences. *Int. Rev. Appl. Linguist. Lang. Teach.* 44, 231–250. doi: 10.1515/IRAL.2006.010
- Bandler, R. J. (1969). Facilitation of aggressive behaviour in rat by direct cholinergic stimulation of the hypothalamus. *Nature* 224, 1035–1036. doi: 10.1038/2241035a0
- Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *J. Mem. Lang.* 68, 255–278. doi: 10.1016/j.jml.2012.11.001
- Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01
- Berthoud, H. R., and Neuhuber, W. L. (2000). Functional and chemical anatomy of the afferent vagal system. *Auton. Neurosci.* 85, 1–17. doi: 10.1016/S1566-0702(00)00215-0
- Best, C. T. (1994). The emergence of native-language phonological influences in infants: a perceptual assimilation model. *Dev. Speech Percept.* 167, 233–277.
- Best, C. T., and Tyler, M. D. (2007). “Nonnative and second-language speech perception: Commonalities and complementarities,” in *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*, eds. M. J. Munro, and O.-S. Bohn (Amsterdam: John Benjamins), 13–34.
- Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., and Tohkura, Y. I. (1997). Training Japanese listeners to identify English /r/ and /l/. IV. Some effects of perceptual learning on speech production. *J. Acoust. Soc. Am.* 101, 2299–2310. doi: 10.1121/1.418276
- Brosseau-Lapr e, F., Rvachew, S., Clayards, M., and Dickson, D. (2013). Stimulus variability and perceptual learning of nonnative vowel categories. *Appl. Psycholinguist.* 34, 419–441. doi: 10.1017/S0142716411000750
- Burger, A. M., Van der Does, W., Thayer, J. F., Brosschot, J. F., and Verkuil, B. (2019). Transcutaneous vagus nerve stimulation reduces spontaneous but not induced negative thought intrusions in high worriers. *Biol. Psychol.* 142, 80–89. doi: 10.1016/j.biopsycho.2019.01.014
- Burger, A. M., Verkuil, B., Fenlon, H., Thijs, L., Cools, L., Miller, H. C., et al. (2017). Mixed evidence for the potential of non-invasive transcutaneous vagal nerve stimulation to improve the extinction and retention of fear. *Behav. Res. Ther.* 97, 64–74. doi: 10.1016/j.brat.2017.07.005
- Butt, M. F., Albusoda, A., Farmer, A. D., and Aziz, Q. (2020). The anatomical basis for transcutaneous auricular vagus nerve stimulation. *J. Anat.* 236, 588–611. doi: 10.1111/joa.13122
- Calloway, R., Karuzis, V., Tseng, A., Martinez, D., and O’Rourke, P. (2020). “Auricular transcutaneous vagus nerve stimulation (tVNS) affects mood and anxiety during second language learning,” in *Proceedings of the Annual Meeting of the Cognitive Science Society*, Vol. 42, 3496–3502.
- Chen, L., Zhang, J., Wang, Z., Zhang, X., Zhang, L., Xu, M., et al. (2021). Effects of transcutaneous vagus nerve stimulation (tVNS) on action planning: a behavioural and EEG study. *IEEE Transact. Neural Syst. Rehabil. Eng.* 30, 1675–1683. doi: 10.1109/TNSRE.2021.3131497
- Chen, Y., Lu, X., and Hu, L. (2023). Transcutaneous auricular vagus nerve stimulation facilitates cortical arousal and alertness. *Int. J. Environ. Res. Public Health* 20:1402. doi: 10.3390/ijerph20021402
- Christensen, R. H. B. (2022). *ordinal - Regression Models for Ordinal Data. R package version 2022.11-16*. Available online at: <https://CRAN.R-project.org/package=Sordinal> (accessed November 16, 2022).
- Clark, K. B., Krahl, S. E., Smith, D. C., and Jensen, R. A. (1995). Post-training unilateral vagal stimulation enhances retention performance in the rat. *Neurobiol. Learn. Mem.* 63, 213–216. doi: 10.1006/nlme.1995.1024
- Clark, K. B., Naritoku, D. K., Smith, D. C., Browning, R. A., and Jensen, R. A. (1999). Enhanced recognition memory following vagus nerve stimulation in human subjects. *Nat. Neurosci.* 2, 94–98. doi: 10.1038/4600
- Coffey, E. B. J., Herholz, S. C., Scala, S., and Zatorre, R. J. (2011). “Montreal Music History Questionnaire: a tool for the assessment of music-related experience in music cognition research,” in *The Neurosciences and Music IV: Learning and Memory, Conference* (Edinburgh).
- Collins, L., Boddington, L., Steffan, P. J., and McCormick, D. (2021). Vagus nerve stimulation induces widespread cortical and behavioral activation. *Curr. Biol.* 31, 2088–2098. doi: 10.1016/j.cub.2021.02.049
- De Ridder, D., Vanneste, S., Engineer, N. D., and Kilgard, M. P. (2014). Safety and efficacy of vagus nerve stimulation paired with tones for the treatment of tinnitus: a case series. *Neuromodulation* 17, 170–179. doi: 10.1111/ner.12127
- Diaz, B., Mitterer, H., Broersma, M., and Sebastián-Gallés, N. (2012). Individual differences in late bilinguals’ L2 phonological processes: from acoustic-phonetic analysis to lexical access. *Learn. Individ. Differ.* 22, 680–689. doi: 10.1016/j.lindif.2012.05.005
- Dörnyei, Z. (2001). New themes and approaches in second language motivation research. *Annu. Rev. Appl. Linguist.* 21, 43–59. doi: 10.1017/S0267190501000034
- Dörnyei, Z. (2009). “The L2 motivational self system,” in *Motivation, Language Identity and the L2 Self*, eds. Z. Dörnyei, and E. Ushioda (Bristol: Multilingual Matters).
- Elger, G., Hoppe, C., Falkai, P., Rush, A. J., and Elger, C. E. (2000). Vagus nerve stimulation is associated with mood improvements in epilepsy patients. *Epilepsy Res.* 42, 203–210. doi: 10.1016/S0920-1211(00)00181-9
- Engineer, C. T., Engineer, N. D., Riley, J. R., Seale, J. D., and Kilgard, M. P. (2015). Pairing speech sounds with vagus nerve stimulation drives stimulus-specific cortical plasticity. *Brain Stimul.* 8, 637–644. doi: 10.1016/j.brs.2015.01.408
- Ferstl, M., Teckentrup, V., Lin, W. M., Kräutlein, F., Kühnel, A., Klaus, J., et al. (2022). Non-invasive vagus nerve stimulation boosts mood recovery after effort exertion. *Psychol. Med.* 52, 3029–3039. doi: 10.1017/S0033291720005073
- Frangos, E., Ellrich, J., and Komisaruk, B. R. (2015). Non-invasive access to the vagus nerve central projections via electrical stimulation of the external ear: fMRI evidence in humans. *Brain Stimul.* 8, 624–636. doi: 10.1016/j.brs.2014.11.018
- Gardner, R. C. (2000). Correlation, causation, motivation, and second language acquisition. *Can. Psychol.* 41:10. doi: 10.1037/h0086854
- Ghacibeh, G. A., Shenker, J. I., Shenal, B., Uthman, B. M., and Heilman, K. M. (2006). The influence of vagus nerve stimulation on memory. *Cogn. Behav. Neurol.* 19, 119–122. doi: 10.1097/01.wnn.0000213908.34278.7d
- Giannakopoulou, A., Brown, H., Clayards, M., and Wonnacott, E. (2017). High or low? Comparing high and low-variability phonetic training in adult and child second language learners. *PeerJ* 5:e3209. doi: 10.7717/peerj.3209
- Gignac, G. E., and Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Pers. Individ. Dif.* 102, 74–78. doi: 10.1016/j.paid.2016.06.069
- Giraudier, M., Ventura-Bort, C., and Weymar, M. (2020). Transcutaneous vagus nerve stimulation (tVNS) improves high-confidence recognition memory but not emotional word processing. *Front. Psychol.* 11:1276. doi: 10.3389/fpsyg.2020.01276
- Gisev, N., Bell, J. S., and Chen, T. F. (2013). Interrater agreement and interrater reliability: key concepts, approaches, and applications. *Res. Soc. Administr. Pharm.* 9, 330–338. doi: 10.1016/j.sapharm.2012.04.004
- GoldWave Inc. (2015). *GoldWave (Version 6.15)* [Computer Software]. Available online at: <https://www.goldwave.ca/>
- Golestani, N., and Zatorre, R. J. (2004). Learning new sounds of speech: reallocation of neural substrates. *Neuroimage* 21, 494–506. doi: 10.1016/j.neuroimage.2003.09.071
- Hallett, M. (2000). Transcranial magnetic stimulation and the human brain. *Nature* 406, 147–150. doi: 10.1038/35018000
- Hanulíková, A., Dediú, D., Fang, Z., Bašnaková, J., and Huettig, F. (2012). Individual differences in the acquisition of a complex L2 phonology: a training study. *Lang. Learn.* 62, 79–109. doi: 10.1111/j.1467-9922.2012.00707.x
- Honda, C. T., Clayards, M., and Baum, S. R. (2024). Exploring individual differences in native phonetic perception and their link to nonnative phonetic perception. *J. Exp. Psychol.* 50, 370–394. doi: 10.1037/xhp0001191
- Hsieh, L., Gandour, J., Wong, D., and Hutchins, G. D. (2001). Functional heterogeneity of inferior frontal gyrus is shaped by linguistic experience. *Brain Lang.* 76, 227–252. doi: 10.1006/brln.2000.2382
- Iverson, P., Hazan, V., and Bannister, K. (2005). Phonetic training with acoustic cue manipulations: a comparison of methods for teaching English /r-/l/ to Japanese adults. *J. Acoust. Soc. Am.* 118, 3267–3278. doi: 10.1121/1.2062307
- Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Kettermann, A., and Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition* 87, B47–B57. doi: 10.1016/S0010-0277(02)00198-1
- Jacobs, H. I., Riphagen, J. M., Razat, C. M., Wiese, S., and Sack, A. T. (2015). Transcutaneous vagus nerve stimulation boosts associative memory in older individuals. *Neurobiol. Aging* 36, 1860–1867. doi: 10.1016/j.neurobiolaging.2015.02.023
- Johnson, J. S., and Newport, E. L. (1989). Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cogn. Psychol.* 21, 60–99. doi: 10.1016/0010-0285(89)90003-0
- Kaan, E., De Aguiar, I., Clarke, C., Lamb, D. G., Williamson, J. B., and Porges, E. C. (2021). A transcutaneous vagus nerve stimulation study on verbal order memory. *J. Neurolinguist.* 59:100990. doi: 10.1016/j.jneuroling.2021.100990
- Kang, K. H., and Guion, S. G. (2006). Phonological systems in bilinguals: age of learning effects on the stop consonant systems of Korean-English bilinguals. *J. Acoust. Soc. Am.* 119, 1672–1683. doi: 10.1121/1.2166607
- Koo, T. K., and Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* 15, 155–163. doi: 10.1016/j.jcm.2016.02.012
- Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. (2017). lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* 82, 1–26. doi: 10.18637/jss.v082.i13
- Lai, J., and David, S. V. (2021). Short-term effects of vagus nerve stimulation on learning and evoked activity in auditory cortex. *eNeuro* 8:ENEURO.0522-20.2021. doi: 10.1523/ENEURO.0522-20.2021

- Landis, J. R., and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics* 33, 159–174. doi: 10.2307/2529310
- Li, P., Zhang, F., Tsai, E., and Puls, B. (2013). Language history questionnaire (lhq 2.0): a new dynamic web-based research tool. *Biling. Lang. Cogn.* 17, 673–680. doi: 10.1017/S1366728913000606
- Liu, A., Rong, P., Gong, L., Song, L., Wang, X., Li, L., et al. (2018). Efficacy and safety of treatment with transcutaneous vagus nerve stimulation in 17 patients with refractory epilepsy evaluated by electroencephalogram, seizure frequency, and quality of life. *Med. Sci. Monit. Int. Med. J. Exp. Clin. Res.* 24, 8439–8448. doi: 10.12659/MSM.910689
- Llanos, F., McHaney, J. R., Schuerman, W. L., Han, G. Y., Leonard, M. K., and Chandrasekaran, B. (2020). Non-invasive peripheral nerve stimulation selectively enhances speech category learning in adults. *NPJ Sci. Learn.* 5, 1–11. doi: 10.1038/s41539-020-0070-0
- Masgoret, A. M., and Gardner, R. C. (2003). Attitudes, motivation, and second language learning: a meta-analysis of studies conducted by Gardner and associates. *Lang. Learn.* 53, 167–210. doi: 10.1111/1467-9922.00227
- Mather, M., Clewett, D., Sakaki, M., and Harley, C. W. (2016). Norepinephrine ignites local hotspots of neuronal excitation: how arousal amplifies selectivity in perception and memory. *Behav. Brain Sci.* 39:667. doi: 10.1017/S0140525X15000667
- Mather, M., and Sutherland, M. R. (2011). Arousal-biased competition in perception and memory. *Perspect. Psychol. Sci.* 6, 114–133. doi: 10.1177/1745691611400234
- Mathworks (2017). *Matlab (Version 2017)* [Computer Software]. Available online at: <https://www.mathworks.com/products/matlab.html>
- Maye, J., Werker, J. F., and Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82, B101–B111. doi: 10.1016/S0010-0277(01)00157-3
- Mayr, R., and Escudero, P. (2015). Explaining individual variation in L2 perception: rounded vowels in English learners of German. *Biling. Lang. Cogn.* 13, 279–297. doi: 10.1017/S1366728909990022
- McAuley, E., Duncan, T., and Tammen, V. V. (1989). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: a confirmatory factor analysis. *Res. Q. Exerc. Sport* 60, 48–58. doi: 10.1080/02701367.1989.10607413
- McCandliss, B. D., Fiez, J. A., Protopapas, A., Conway, M., and McClelland, J. L. (2002). Success and failure in teaching the [r]-[l] contrast to Japanese adults: tests of a Hebbian model of plasticity and stabilization in spoken language perception. *Cogn. Affect. Behav. Neurosci.* 2, 89–108. doi: 10.3758/CABN.2.2.89
- McHaney, J. R., Schuerman, W. L., Leonard, M. K., and Chandrasekaran, B. (2023). Transcutaneous auricular vagus nerve stimulation modulates performance but not pupil size during nonnative speech category learning. *J. Speech Lang. Hear. Res.* 66, 3825–3843. doi: 10.1044/2023_JSLHR-22-00596
- Moulton, W. G. (1962). *The Sounds of English and German*. Chicago, IL: University of Chicago Press.
- Munro, M., and Mann, V. (2005). Age of immersion as a predictor of foreign accent. *Appl. Psychol.* 26, 311–341. doi: 10.1017/S0142716405050198
- Neuser, M. P., Teckentrup, V., Kühnel, A., Hallschmid, M., Walter, M., and Kroemer, N. B. (2020). Vagus nerve stimulation boosts the drive to work for rewards. *Nat. Commun.* 11:3555. doi: 10.1038/s41467-020-17344-9
- Pandža, N. B., Phillips, I., Karuzis, V. P., O'Rourke, P., and Kuchinsky, S. E. (2020). Neurostimulation and pupillometry: new directions for learning and research in applied linguistics. *Annu. Rev. Appl. Linguist.* 40, 56–77. doi: 10.1017/S0267190520000069
- Perfors, A., and Ong, J. (2012). “Musicians are better at learning non-native sound contrasts even in non-tonal languages,” in *Proceedings of the Annual Meeting of the Cognitive Science Society, Vol. 34*, 839–844.
- Phillips, I., Calloway, R. C., Karuzis, V. P., Pandža, N. B., O'Rourke, P., and Kuchinsky, S. E. (2021). Transcutaneous auricular vagus nerve stimulation strengthens semantic representations of foreign language tone words during initial stages of learning. *J. Cogn. Neurosci.* 34, 127–152. doi: 10.1162/jocn_a_01783
- Plonsky, L., and Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Lang. Learn.* 64, 878–912. doi: 10.1111/lang.12079
- Pulvermüller, F., and Schumann, J. H. (1994). Neurobiological mechanisms of language acquisition. *Lang. Learn.* 44, 681–734. doi: 10.1111/j.1467-1770.1994.tb00635.x
- R Development Core Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online at: <https://www.R-project.org/> (accessed June 16, 2021).
- Reetzke, R., Xie, Z., Llanos, F., and Chandrasekaran, B. (2018). Tracing the trajectory of sensory plasticity across different stages of speech learning in adulthood. *Curr. Biol.* 28, 1419–1427. doi: 10.1016/j.cub.2018.03.026
- Rizzo, P., Beelke, M., De Carli, F., Canovaro, P., Nobili, L., Robert, A., et al. (2003). Chronic vagus nerve stimulation improves alertness and reduces rapid eye movement sleep in patients affected by refractory epilepsy. *Sleep* 26, 607–611. doi: 10.1093/sleep/26.5.607
- Rufener, K. S., Geyer, U., Janitzky, K., Heinze, H. J., and Zaehle, T. (2018). Modulating auditory selective attention by non-invasive brain stimulation: differential effects of transcutaneous vagal nerve stimulation and transcranial random noise stimulation. *Eur. J. Neurosci.* 48, 2301–2309. doi: 10.1111/ejn.14128
- Rufener, K. S., Wienke, C., Salanje, A., Haghikia, A., and Zaehle, T. (2023). Effects of transcutaneous auricular vagus nerve stimulation paired with tones on electrophysiological markers of auditory perception. *Brain Stimul.* 16, 982–989. doi: 10.1016/j.brs.2023.06.006
- Saito, H. (2021). *The Effects of Frequency and Type of Feedback on L2 Phoneme Learning* (Doctoral dissertation). Montreal: McGill University.
- Saito, K., Dewaele, J. M., Abe, M., and In'nami, Y. (2018). Motivation, emotion, learning experience, and second language comprehensibility development in classroom settings: a cross-sectional and longitudinal study. *Lang. Learn.* 68, 709–743. doi: 10.1111/lang.12297
- Sakai, M., and Moorman, C. (2018). Can perception training improve the production of second language phonemes? A meta-analytic review of 25 years of perception training research. *Appl. Psycholinguist.* 39, 187–224. doi: 10.1017/S0142716417000418
- Sawchenko, P. E. (1983). Central connections of the sensory and motor nuclei of the vagus nerve. *J. Auton. Nerv. Syst.* 9, 13–26. doi: 10.1016/0165-1838(83)90129-7
- Schuerman, W. L., Chandrasekaran, B., and Leonard, M. K. (2022). Arousal states as a key source of variability in speech perception and learning. *Languages* 7:19. doi: 10.3390/languages7010019
- Sellaro, R., van Leusden, J. W., Tona, K. D., Verkuil, B., Nieuwenhuis, S., and Colzato, L. S. (2015). Transcutaneous vagus nerve stimulation enhances post-error slowing. *J. Cogn. Neurosci.* 27, 2126–2132. doi: 10.1162/jocn_a_00851
- Sharon, O., Fahoum, F., and Nir, Y. (2021). Transcutaneous vagus nerve stimulation in humans induces pupil dilation and attenuates alpha oscillations. *J. Neurosci.* 41, 320–330. doi: 10.1523/JNEUROSCI.1361-20.2020
- Shetake, J. A., Engineer, N. D., Vrana, W. A., Wolf, J. T., and Kilgard, M. P. (2012). Pairing tone trains with vagus nerve stimulation induces temporal plasticity in auditory cortex. *Exp. Neurol.* 233, 342–349. doi: 10.1016/j.expneurol.2011.10.026
- Shim, H. J., Kwak, M. Y., An, Y. H., Kim, D. H., Kim, Y. J., and Kim, H. J. (2015). Feasibility and safety of transcutaneous vagus nerve stimulation paired with notched music therapy for the treatment of chronic tinnitus. *J. Audiol. Otol.* 19:159. doi: 10.7874/jao.2015.19.3.159
- Slevc, L. R., and Miyake, A. (2006). Individual differences in second-language proficiency: does musical ability matter?. *Psychol. Sci.* 17, 675–681. doi: 10.1111/j.1467-9280.2006.01765.x
- So, C. K., and Best, C. T. (2014). Phonetic influences on English and French listeners' assimilation of Mandarin tones to native prosodic categories. *Studi. Sec. Lang. Acquisit.* 36, 195–221. doi: 10.1017/S0272263114000047
- Stowe, L. A., and Sabourin, L. (2005). Imaging the processing of a second language: effects of maturation and proficiency on the neural processes involved. *Int. Rev. Appl. Linguist. Lang. Teach.* 43, 329–353. doi: 10.1515/iral.2005.43.4.329
- Strange, W., Bohn, O. S., Trent, S. A., and Nishi, K. (2004). Acoustic and perceptual similarity of North German and American English vowels. *J. Acoust. Soc. Am.* 115, 1791–1807. doi: 10.1121/1.1687832
- Strange, W., and Dittmann, S. (1984). Effects of discrimination training on the perception of /r-/l/ by Japanese adults learning English. *Percept. Psychophys.* 36, 131–145. doi: 10.3758/BF03202673
- Sun, J. B., Cheng, C., Tian, Q. Q., Yuan, H., Yang, X. J., Deng, H., et al. (2021). Transcutaneous auricular vagus nerve stimulation improves spatial working memory in healthy young adults. *Front. Neurosci.* 15:790793. doi: 10.3389/fnins.2021.790793
- Sun, L., Peräkylä, J., Holm, K., Haapasalo, J., Lehtimäki, K., Ogawa, K. H., et al. (2017). Vagus nerve stimulation improves working memory performance. *J. Clin. Exp. Neuropsychol.* 39, 954–964. doi: 10.1080/13803395.2017.1285869
- Thair, H., Holloway, A. L., Newport, R., and Smith, A. D. (2017). Transcranial direct current stimulation (tDCS): a beginner's guide for design and implementation. *Front. Neurosci.* 11:276151. doi: 10.3389/fnins.2017.00641
- Thakkar, V. J., Engelhart, A. S., Khodaparast, N., Abadzi, H., and Centanni, T. M. (2020). Transcutaneous auricular vagus nerve stimulation enhances learning of novel letter-sound relationships in adults. *Brain Stimul.* 13, 1813–1820. doi: 10.1016/j.brs.2020.10.012
- Thakkar, V. J., Richardson, Z. A., Dang, A., and Centanni, T. M. (2023). The effect of non-invasive vagus nerve stimulation on memory recall in reading: a pilot study. *Behav. Brain Res.* 438:114164. doi: 10.1016/j.bbr.2022.114164
- Van Leusden, J. W., Sellaro, R., and Colzato, L. S. (2015). Transcutaneous Vagal Nerve Stimulation (tVNS): a new neuromodulation tool in healthy humans? *Front. Psychol.* 6:102. doi: 10.3389/fpsyg.2015.00102
- Ventura-Bort, C., Wirkner, J., Genheimer, H., Wendt, J., Hamm, A. O., and Weymar, M. (2018). Effects of transcutaneous vagus nerve stimulation (tVNS) on

the P300 and alpha-amylase level: a pilot study. *Front. Hum. Neurosci.* 12:202. doi: 10.3389/fnhum.2018.00202

Villani, V., Tsakiris, M., and Azevedo, R. T. (2019). Transcutaneous vagus nerve stimulation improves interoceptive accuracy. *Neuropsychologia* 134:107201. doi: 10.1016/j.neuropsychologia.2019.107201

Weber-Fox, C. M., and Neville, H. J. (1996). Maturation constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *J. Cogn. Neurosci.* 8, 231–256. doi: 10.1162/jocn.1996.8.3.231

White, E. J., Hutka, S. A., Williams, L. J., and Moreno, S. (2013). Learning, neural plasticity and sensitive periods: implications for language acquisition, music training and transfer across the lifespan. *Front. Syst. Neurosci.* 7:90. doi: 10.3389/fnsys.2013.00090

Williams, M., Burden, R., Poulet, G., and Maun, I. (2004). Learners' perceptions of their successes and failures in foreign language learning. *Lang. Learn. J.* 30, 19–29. doi: 10.1080/09571730485200191

Yakunina, N., Kim, S. S., and Nam, E. C. (2017). Optimization of transcutaneous vagus nerve stimulation using functional MRI. *Neuromodulation* 20, 290–300. doi: 10.1111/ner.12541

Yuan, H., and Silberstein, S. D. (2016a). Vagus nerve and vagus nerve stimulation, a comprehensive review: part I. *Headache* 56, 71–78. doi: 10.1111/head.12647

Yuan, H., and Silberstein, S. D. (2016b). Vagus nerve and vagus nerve stimulation, a comprehensive review: part II. *Headache* 56, 259–266. doi: 10.1111/head.12650

Zuo, Y., Smith, D. C., and Jensen, R. A. (2007). Vagus nerve stimulation potentiates hippocampal LTP in freely-moving rats. *Physiol. Behav.* 90, 583–589. doi: 10.1016/j.physbeh.2006.11.009