

Review

Arousal States as a Key Source of Variability in Speech Perception and Learning

William L. Schuerman ¹, Bharath Chandrasekaran ² and Matthew K. Leonard ^{1,*}

¹ Department of Neurological Surgery, Weill Institute for Neurosciences, University of California, San Francisco, CA 94158, USA; william.schuerman@ucsf.edu

² Center for the Neural Basis of Cognition, Department of Communication Science and Disorders, School of Health and Rehabilitation Sciences, University of Pittsburgh, Pittsburgh, PA 15213, USA; b.chandra@pitt.edu

* Correspondence: matthew.leonard@ucsf.edu

Abstract: The human brain exhibits the remarkable ability to categorize speech sounds into distinct, meaningful percepts, even in challenging tasks like learning non-native speech categories in adulthood and hearing speech in noisy listening conditions. In these scenarios, there is substantial variability in perception and behavior, both across individual listeners and individual trials. While there has been extensive work characterizing stimulus-related and contextual factors that contribute to variability, recent advances in neuroscience are beginning to shed light on another potential source of variability that has not been explored in speech processing. Specifically, there are task-independent, moment-to-moment variations in neural activity in broadly-distributed cortical and subcortical networks that affect how a stimulus is perceived on a trial-by-trial basis. In this review, we discuss factors that affect speech sound learning and moment-to-moment variability in perception, particularly arousal states—neurotransmitter-dependent modulations of cortical activity. We propose that a more complete model of speech perception and learning should incorporate subcortically-mediated arousal states that alter behavior in ways that are distinct from, yet complementary to, top-down cognitive modulations. Finally, we discuss a novel neuromodulation technique, transcutaneous auricular vagus nerve stimulation (taVNS), which is particularly well-suited to investigating causal relationships between arousal mechanisms and performance in a variety of perceptual tasks. Together, these approaches provide novel testable hypotheses for explaining variability in classically challenging tasks, including non-native speech sound learning.

Keywords: second language acquisition; speech perception; arousal; neuromodulation; pupillometry; transcutaneous auricular vagus nerve stimulation



Citation: Schuerman, William L., Bharath Chandrasekaran, and Matthew K. Leonard. 2022. Arousal States as a Key Source of Variability in Speech Perception and Learning. *Languages* 7: 19. <https://doi.org/10.3390/languages7010019>

Academic Editor: David P Birdsong

Received: 7 October 2021

Accepted: 19 January 2022

Published: 24 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The ability to perceive speech, especially under challenging conditions, reflects a remarkable set of computational processes that the human brain is well-adapted to perform. A major challenge that both expert and novice listeners face when learning a new language is substantial variability in input across different speakers, contexts, and listening environments. To comprehend spoken language, listeners must transform a highly variable and often noisy acoustic signal into meaningful linguistic units. Listeners use numerous sources of knowledge to overcome this variability, including (but not limited to) visual information (Campbell 2008; McGurk and MacDonald 1976), coarticulation (Kang et al. 2016; Mann and Repp 1980), lexical status (Ganong 1980; Luthra et al. 2021), semantic information (Kutas and Federmeier 2000; Miller and Isard 1963), and discourse structure (Brouwer et al. 2013; Van Berkum et al. 2005). In addition, the active processes of perception and comprehension are modulated by task- and goal-driven factors like attention (Heald and Nusbaum 2014; Huyck and Johnsrude 2012). Together, these factors provide listeners with the flexibility necessary to comprehend speech, including under conditions where the signal-to-noise

ratio of the input is decreased (Guediche et al. 2014), or in the context of challenging tasks like acquisition of an unfamiliar language (Birdsong 2018).

However, the same speech sounds can be perceived quite differently due to ambiguity in the input (e.g., masking in a noisy environment) or ambiguity in the listener's perceptual or cognitive representations of speech (e.g., due to variation in familiarity with the language). While there has been extensive work characterizing the contributions of these stimulus-related factors to trial-by-trial variability in perception (Guediche et al. 2014; Heald and Nusbaum 2014), recent advances in neuroscience are beginning to shed light on another, complementary set of factors that may be just as important to understanding behavioral variability. Specifically, moment-to-moment variations in neural activity that are not directly related to a given task may play a key role in perceptual and behavioral outcomes.

In this review, we examine the evidence for the ability of **arousal states**—neurotransmitter-dependent modulations of cortical activity—to affect behaviors that are central to the ability to learn new languages. Specifically, we focus on how arousal states may influence non-native sound learning in adulthood and perception of acoustically ambiguous speech, since these constitute concrete examples of core processes that reflect the tight coordination between cortical perceptual systems and subcortical arousal systems. While arousal states are also likely to be involved in learning to produce the sounds of a non-native language, here we refer to “sound learning” solely with regard to the process of learning to identify and discriminate non-native speech sound categories.

First, we provide an overview of the physiological basis of arousal states and describe how brain activity may be modulated by fluctuations in these systems. Second, we will discuss the challenging task of learning to discriminate and identify non-native speech sound categories in adulthood, focusing in particular on our current understanding of the neural processes that the arousal system may act upon. Third, we will discuss examples of moment-to-moment perceptual variability that provide crucial insight into factors that affect how speech sounds are processed. Fourth, we propose that changes in arousal states may be able to explain this moment-to-moment variability. Finally, we introduce an emerging neuromodulation technique, non-invasive transcutaneous auricular vagus nerve stimulation (taVNS), that can be used to provide causal tests of arousal mechanisms in a variety of tasks, including sound learning. We propose that taVNS holds promise as both a scientific and translational tool for understanding and manipulating arousal states that may play a major role in learning and perceptual outcomes.

2. The Physiology of Arousal States

The arousal system is one of the most fundamental mechanisms in the vertebrate central nervous system (Coull 1998; Whyte 1992). It is composed of multiple overlapping components that modulate core bodily functions and states including wakefulness/alertness, body-wide motoric activity, and affective reactivity (Satpute et al. 2019), as well as other functions such as neural plasticity (Coull et al. 1999; Martins and Froemke 2015; Unsworth and Robison 2017).

Arousal is generally divided into two subtypes with related but dissociable mechanisms. Tonic arousal is characterized by slow fluctuations, such as those governed by circadian rhythm, and corresponds behaviorally to states of wakefulness/alertness. In contrast, phasic arousal refers to rapid (i.e., on the level of seconds and milliseconds) fluctuations in neural responsivity operating within stages of tonic arousal (Satpute et al. 2019; Whyte 1992). While tonic arousal may have important implications for perceptual processes, moment-to-moment variability in perception is more strongly influenced by phasic arousal (McGinley et al. 2015b).

Anatomically, the arousal system consists of numerous ascending and descending pathways originating in the lower brainstem. These pathways activate subcortical structures that release neurotransmitters such as norepinephrine (NE) and acetylcholine (ACh; Quinkert et al. 2011). The release of these neurotransmitters modulates activity across the cortex. Among these pathways, the locus coeruleus-norepinephrine (LC-NE) pathway is

one of the most important systems for controlling phasic arousal (Aston-Jones and Cohen 2005; Sara 2009; Unsworth and Robison 2017). The LC is a brainstem nucleus that is the main source of NE and projects to widespread cortical and subcortical sites (Berridge and Waterhouse 2003; McCormick and Pape 1990; Ranjbar-Slamloo and Fazlali 2020). In addition to controlling arousal, activation of the LC has been found to affect numerous cognitive processes such as attention, memory, and sensory processing (Poe et al. 2020; Sara and Bouret 2012). Crucially, the LC-NE system has rapid effects on cortical activity, reflected in dynamic changes in neurophysiology and psychophysiological measures like pupil dilation (Gilzenrat et al. 2010; McGinley et al. 2015b; Reimer et al. 2016).

Levels of arousal are often divided into distinct states, corresponding to low, moderate, and high arousal (Aston-Jones and Cohen 2005). While arousal states have long been appreciated in many other domains (e.g., Symmes and Anderson 1967), to date, there is little work on their role in speech perception and learning. In general, arousal states can be viewed as varying degrees of receptivity (how likely it is that a population will be activated by some input) and reactivity (how strongly a population responds to a given input) in neural populations. Brief changes in phasic arousal induced via air puff strongly suppress auditory responses in avian sensorimotor regions (Cardin and Schmidt 2003), and direct application of NE to auditory processing areas enhances or suppresses auditory responses dependent on dosage (Cardin and Schmidt 2004). In mouse models, increased arousal broadens the response bandwidth of cells tuned to specific frequencies (Lin et al. 2019) which may affect sensory processing by generating a larger neural response to a particular stimulus. Similarly, the performance of mice attempting to detect a pure tone inside a complex auditory mask has been shown to be modulated by moment-to-moment arousal state (McGinley et al. 2015a).

Emerging research in many domains, including working memory, attentional control (Unsworth and Robison 2017), and plasticity (Martins and Froemke 2015), collectively points to the importance of neuromodulator-dependent arousal states as a critical source of individual neural variability. With regards to speech, prior studies have tended to focus on how arousal states affect production (e.g., Kleinow and Smith 2006) or how perception and comprehension may affect arousal (e.g., Zekveld et al. 2018). Research that focuses specifically on the effect of arousal states (independent of anxiety, c.f. Mattys et al. 2013) on perception and learning is only recently beginning to emerge. For example, a recent EEG study found that during sleep, neural responses phase locked to speech (isolated vowel sounds) are modulated by changes in arousal state (Mai et al. 2019). Similarly, Llanos et al. (2020) recently demonstrated that performance on a non-native speech sound learning could be modulated using stimulation techniques that target arousal mechanisms. We propose that these rapid fluctuations in arousal states may explain moment-to-moment variability in neural activity and behavior during speech perception, with strong implications for explaining variability in sound learning in adulthood.

3. Emergence of Non-Native Speech Category Representations in Adulthood

A clear example of variability in speech perception comes from the domain of non-native speech sound learning. Difficulties with perceptual discrimination and identification can disproportionately affect progress when learning a new language. After developmental sensitive periods (usually 4–12 months; Kral et al. 2019; Kuhl 2010, 2004), there is greater neural commitment (and preference) to the phonological inventories of languages that the infant has been exposed to and reduced neural commitment (and preference) to non-native sound categories (Werker and Hensch 2015; Yu and Zhang 2018). This early experience can impact later learning (Finn et al. 2013; Kuhl et al. 2005; though c.f. Birdsong and Vanhove 2016), making it notoriously difficult for adults to acquire certain non-native speech sound categories (e.g., Japanese learners acquiring /r/ vs. /l/ distinctions; Bradlow 2008; Zhang et al. 2005). However, laboratory-based training approaches have demonstrated that robust and generalizable learning is achievable in adulthood and that this learning is retained well after the training period (Myers 2014; Reetzke et al. 2018). Training approaches widely

differ, as do the learning outcomes across individuals (Chandrasekaran et al. 2010; Golestani and Zatorre 2009). Some approaches use synthesized stimuli with acoustic information constrained to generate specific contrasts (Scharinger et al. 2013) or reflect native-like distributions (Reetzke et al. 2018; Zhang et al. 2009), while others use naturalistic stimuli without distributional constraints (Chandrasekaran et al. 2010; Sadakata and McQueen 2013). Some actively leverage talker variability (Brosseau-Lapr   et al. 2013; Perrachione et al. 2011), others do not. Some approaches involve no feedback, some involve incidental or implicit feedback (Lim and Holt 2011), while some provide varying levels of explicit feedback (Yi and Chandrasekaran 2016).

Despite the large number of training approaches, two consistent themes emerge from the speech sound learning literature: First, on average, adults can learn even difficult non-native phonetic categories, and second, large-scale individual differences persist across various training approaches. However, it remains unclear what constitutes the underlying neural mechanisms and sources of individual variability.

A majority of laboratory-based phonetic training approaches have used three training characteristics: (1) naturalistic stimuli produced by native speakers, (2) high-variability—using multiple talkers and segmental contexts, and (3) trial-by-trial feedback. While #1 and #2 help listeners focus on dimensions that are category-relevant (and ignore dimensions that are highly variable across talkers/segments), #3 allows listeners the opportunity to monitor and learn from errors. These three training characteristics have engendered significant and generalizable learning for various speech categories (e.g., the /l/~/r/ contrast, lexical tone, voice-onset-time distinctions, etc.). Other approaches (e.g., incidental learning and implicit learning) also result in significant and generalizable learning. While it is important to note that other approaches (e.g., incidental and implicit learning) also have been shown to improve performance in language learning tasks, in this review we restrict our discussions to training approaches with the three training characteristics (natural speech, talker variability, and trial-by-trial feedback) described above.

Over the last several years, a series of studies have adopted a systems neuroscience perspective using multiple neuroimaging methods to provide a better understanding of how sensory and perceptual representations of categories emerge as a function of sound-to-category training. Non-invasive methods like functional magnetic resonance imaging (fMRI) offer high spatial precision and an opportunity to discern network-level activity, while electroencephalography (EEG) offers millisecond precision and an opportunity to discern sensory changes as a function of training (Menon and Crottaz-Herbette 2005). In addition, novel computational techniques such as multivariate decoding and functional connectivity analysis have allowed for a more mechanistic understanding of the neural correlates of speech sound learning.

For example, Feng et al. (2019) examined blood oxygen level-dependent (BOLD) activity using fMRI as participants acquired non-native Mandarin tone categories over a session of sound-to-category training. This study examined how tone category representations emerge within the auditory association cortex across the timescale of a single session (Figure 1). Within a few hundred trials, activation patterns that differentiate tonal categories (syllables produced by different talkers that vary on the basis of pitch patterns) emerged within the left superior temporal gyrus (STG). These emergent representations were robust to talker and segmental variability, suggesting that they reflect abstract category representations. Furthermore, contrasting correct versus incorrect trials showed activation of several striatal regions including the bilateral putamen, caudate nucleus, and nucleus accumbens. Participants who showed increased putamen activation showed more robust learning and employed more procedural-based (sound-to-reward mapping) learning strategies (Chandrasekaran et al. 2014; Maddox and Chandrasekaran 2014).

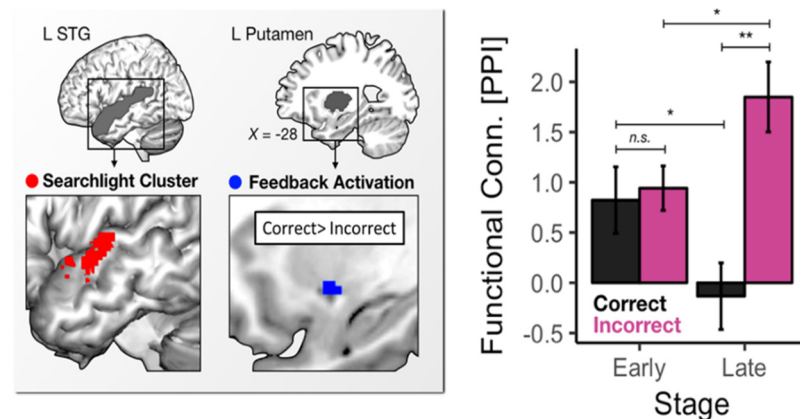


Figure 1. The emergence of neural representations for non-native speech sound categories. Variability-resistant category representations emerge within the left Superior Temporal Gyrus (STG) within a few hundred training trials. Corrective feedback engages the putamen throughout learning. These two regions are more functionally coupled during incorrect feedback (towards the latter half of training) relative to correct feedback. * and ** indicate increasing levels of statistical significance in planned t-tests. Adapted with permission from Feng et al. (2019). Copyright 2019 The Authors.

These results demonstrate that abstract category representations emerge primarily in secondary auditory cortical regions, and that another subcortical network is sensitive to the type of feedback participants receive across trials during learning. Feng et al. (2019) then tested how these two networks interact and found that emergent representations in the left STG were more functionally coupled with the putamen in the latter half of the training paradigm when participants encountered incorrect feedback. They posit that this functional coupling serves to tune emergent category representations in a feedback-dependent manner to ensure continued reward (in the form of correct feedback).

Recently, non-invasive methods like fMRI and EEG have been complemented by invasive direct neural recordings using electrodes implanted on the brain surface. For example, some individuals with conditions such as drug-resistant epilepsy undergo invasive electrophysiological monitoring as part of their clinical care, during which they may elect to participate in research, creating the opportunity to record electrical activity from strips or grids of electrodes placed directly on the surface of the cortex (Chang 2015). These electrocorticographic (ECoG) recordings have high spatial and temporal resolution (with a downside of limited spatial sampling of the brain), making it possible to record the activity of populations of neurons with millisecond precision and with sufficient signal-to-noise to examine activity patterns on a single trial basis and link activity to behavior. Combined with methodological advances that make it possible to track fine-grained neural computations at the level of single individuals and single trials, these approaches have vastly increased our ability to investigate variability in speech perception and sound learning.

A recent ECoG study examined variability in the early stages of non-native speech sound learning (Yi et al. 2021). Native English speakers were trained to identify sounds in Mandarin Chinese, which are distinguished by four distinct pitch patterns (lexical tones) that are difficult for native English speakers to learn (Chandrasekaran et al. 2010). As the participants listened to Mandarin lexical tones and received feedback on their ability to label the tone, a distributed set of neural populations in superior temporal and lateral frontal cortex showed a diverse set of changes, including both increases and decreases in tone-specific encoding patterns depending on whether each trial was behaviorally correct. Both behavior and neural activity were variable across trials (despite participants being presented with the same stimuli), providing a clear neural correlate of trial-wise variability during learning. However, it remains unclear why learning behavior was so variable across trials, in particular leading to non-monotonic learning curves.

While most individuals can acquire novel speech sound categories in adulthood, there are large scale differences in the extent of learning success (e.g., Llanos et al. 2020). Such variability may be driven by individual differences in the extent of engagement of the cortico-striatal network (Yi et al. 2016) and consequently, the robustness of the emergent representations, as well as factors related to the stimuli and training paradigm. Thus, it is crucial to understand the role of moment-to-moment perceptual and behavioral variability in mediating sound learning and more broadly, perception.

4. Moment-to-Moment Variability in Speech Perception and Acquisition

The processes involved in mapping continuous sounds to abstract representations support the goal of successful comprehension. However, these processes also introduce additional sources of variability. For example, the exact same input may be perceived differently when presented multiple times depending on the context (e.g., “bank” as a place to store money versus the edge of a river) and the listener’s prior knowledge (e.g., whether they are familiar with the role of suprasegmental pitch patterns in a tonal language like Mandarin Chinese).

The additional variability introduced by factors associated with the stimulus (e.g., lexical status) or task (e.g., attention, feedback) can explain some but not all trial-to-trial differences in perception and comprehension. Even when context, task, and prior knowledge are held constant, perception can still vary, indicating that there are stimulus- and task-independent processes that strongly influence behavior. Here, we propose another source of variability that explains the apparent non-deterministic nature of perception and learning: phasic fluctuations in arousal states, which modulate brain activity on a moment-by-moment basis, independent of stimulus and task constraints. Despite numerous advances in our understanding of the myriad extrinsic factors that generate variability in speech perception, current models have yet to address these stimulus- and task-independent sources of variability and their role in perception and learning.

In particular, we focus on studies that examine ambiguous sounds that can be perceived in a multistable fashion. Similar to the case of non-native speech sound category learning, these paradigms allow us to study how physically identical sounds are represented in the brain when they are perceived differently.

4.1. Behavioral Evidence for Stimulus- and Task-Independent Variability

One clear demonstration of the behavioral consequences of stimulus- and task-independent variability is the classic psycholinguistic phenomenon of phoneme restoration (Samuel 1981; Warren 1970). When a portion of the acoustic input is completely masked by noise (e.g., the /s/ sound in the word “legislature”), listeners often fail to report that any sound was missing and are unable to localize the noise even when told explicitly that it was there. This effect has been taken as strong evidence for the role of top-down modulation in speech perception; phonological, lexical, and semantic information act to “restore” the missing sound.

However, the restored phoneme that is perceived can change on a trial-by-trial basis. When the noise replaces a phoneme in a word that generates two possible English words (e.g., /fæ#tə/ could be “faster” or “factor”), listeners report bistable perception of the same ambiguous acoustic input. Strikingly, even when provided with strong extrinsic cues (e.g., “She drove the car /fæ#tə/”), perception still varies (Leonard et al. 2016), suggesting that there is an additional source of variability that overrides stimulus characteristics and task goals.

Another common situation in which the same stimulus is repeatedly presented, and when variability in the perceptual response of the listener is a crucial, desired feature, is during second language acquisition. Listeners swiftly adapt to accented speech (Bradlow and Bent 2008; Norris et al. 2003) and with repeated exposure or training can learn to distinguish sounds that were initially indiscriminable (Bradlow 2008). The input may be physically unambiguous, yet perceptual or linguistic representations are less robust,

generating increased variability across trials and among individuals (Chandrasekaran et al. 2010; Paulon et al. 2020; Reetzke et al. 2018). Indeed, decreasing this variability is a primary goal of the learning process, meaning that it is crucial to understand the neural mechanisms that contribute to it.

4.2. Neural Evidence for Arousal-Related Variability

As with speech sound learning (Yi et al. 2021), invasive electrophysiological methods are enabling us to examine the moment-to-moment neural correlates of perceptual variability. A recent ECoG study examined the neural encoding of ambiguous speech sounds in a phoneme restoration task, where trial-by-trial perceptual variability was a key characteristic of the behavior (Leonard et al. 2016). When ECoG participants were presented with both unambiguous (e.g., ‘factor’ and ‘faster’) and ambiguous ([fæ#tə]) stimuli, activity in the superior temporal gyrus (STG) reflected the acoustic-phonetic properties (Mesgarani et al. 2014; Yi et al. 2019) of the perceived sound on a trial-by-trial basis. That is, on a trial when a participant reported hearing the noise burst as /s/ ([fæstə]), neural populations in STG exhibited activity that closely resembled the evoked activity to a real /s/, and the converse was true when the noise was perceived as /k/ ([fæktə]). This activity reflected the online perceptual experience on a trial-by-trial basis, rapidly changing when participants had distinct percepts on repeated presentations of the same stimulus. Strikingly, participants showed semi-random perceptual changes across trials, which was partially explained by activity in a left frontal network. However, it remains unclear what this activity reflects, and importantly, why certain trials were perceived as one word rather than another.

The phenomenon of phoneme restoration represents a clear case in which the stimulus and context are held constant, yet perceptual behavior and its neural correlates are still subject to variation. This leads to a crucial question: What are the sources of this trial-to-trial variability and how do they affect speech perception and speech sound learning on a mechanistic level?

4.3. Cortical State-Dependent Perception and Behavior

Perhaps the most common explanation for trial-by-trial variability in tasks like ambiguous stimulus perception and speech category learning is the interaction between bottom-up perceptual and top-down cognitive processes (Heald and Nusbaum 2014). One source of task-dependent, top-down modulation is a set of fronto-parietal regions known as the Multiple Demand (MD; Duncan 2010; Hasson et al. 2018) network, which is characterized by its involvement in numerous aspects of cognition. The MD network is generally defined as activity that scales with cognitive flexibility, task demands, and engagement of fluid intelligence. The MD network is involved in selecting or generating a response to particular cues (e.g., sensory stimuli) presented in particular contexts. This process is reflected in dynamic reconfigurations of activity patterns across the cortex, resulting in a greater proportion of neurons responding to relevant stimuli and increased similarity in response patterns to targets versus non-targets (Duncan 2010). While there is some debate regarding the role of the MD network in language processing (Diachek et al. 2020), there is clear evidence for its interactive and modulatory effects in trial-to-trial decisions for speech sound categorization, a core behavior for perception and learning. For example, when native listeners must recognize highly variable auditory input as corresponding to a single class (e.g., multiple instances of the same phoneme category spoken by different speakers), categorization is reflected in activity patterns within the MD network (Feng et al. 2021). Interactions between these MD nodes and core speech regions (bilateral STG) were associated with accumulation of evidence for a response, indicating that perceptual behavior is reflected not only in local populations tuned to specific features but also in the coordinated activity of distributed cortical networks.

Outside the MD network, there are other task-dependent processes that can change perceptual and neural representations of sound on a trial-by-trial basis. For example, focusing attention on a particular speaker in a noisy acoustic environment containing competing

signals (e.g., one talker's voice in a room full of other talkers) has been found to enhance or inhibit responses of specific neural populations depending on which stream is being attended to (Brodbeck et al. 2020; Ding and Simon 2012; Mesgarani and Chang 2012). The deployment of attentional resources has been of particular interest to explaining variability with regards to speech perception under adverse listening conditions (for detailed reviews, see Guediche et al. 2014; Scott and McGettigan 2013). For example, momentary fluctuations in attention (indexed by prestimulus alpha phase) have been found to distinguish correct and incorrect lexical decisions (Strauss et al. 2015). Similarly, knowledge about the nature and structure of an unfamiliar stimulus can dramatically change both perception and neural encoding. Noise-vocoded speech (Davis et al. 2005; Davis and Johnsrude 2007) and sine-wave speech (Remez et al. 1981) constitute extreme examples of degraded auditory stimuli that most often are at first completely unintelligible to naïve listeners. However, after an extremely brief (sometimes a single exposure) presentation of information that cues listeners to the nature of these unfamiliar sounds, comprehension is greatly enhanced (Sohoglu et al. 2012), while activity in cortical speech networks suddenly shifts to resemble perception of normal, unfiltered speech (Holdgraf et al. 2016; Khoshkhoo et al. 2018). Adaptation to degraded speech has also been found to recruit executive networks involved in attention and perceptual learning (Erb et al. 2013).

In the case of sound learning, behavior and neural responses have been shown to be affected by task-dependent factors such as instructions or whether a stimulus is presented within a homogenous or diverse set of stimuli. Different types of corrective feedback have been shown to have different effects on trial-to-trial performance when English speakers are learning to identify Mandarin tone speech sounds (Chandrasekaran et al. 2016). When listeners learn to categorize non-native speech sounds presented by multiple talkers, talker-independent representations of those speech categories emerge in STG activity (Feng et al. 2019).

While the preceding examples illustrate cases where task-dependent processes modulate perception and behavior on a trial-by-trial basis, there is still substantial unexplained variability both within and across individuals. For example, variability in the firing patterns of neurons tuned to specific features has been shown to be an important component of the auditory processing system (Faisal et al. 2008; Stein et al. 2005). Even in simple actions such as a button press (Fox et al. 2007), task-independent fluctuations in the activity of cortical regions and networks may account for a considerable proportion of the variance (Sadaghiani and Kleinschmidt 2013; Taghia et al. 2018; Vidaurre et al. 2017). In the case of speech perception, moment-to-moment changes in activity patterns have been found to predict perceptual behavior when input is ambiguous (Leonard et al. 2016). A major unanswered question is what modulates task-independent variability from trial to trial.

We propose that these cortical activity patterns that reflect moment-to-moment variability in perception are an example of **brain states**. Brain states are network-level activity patterns that do not directly encode information like stimulus content, but which dynamically alter stimulus or behavioral representations via changes in functional connectivity (Taghia et al. 2018). Brain states have been well characterized in the domains of sleep (Steriade et al. 2001), memory, and attention (Harris and Thiele 2011), but they have yet to be studied in speech and language. We currently lack an understanding of how brain states modulate representations for speech sounds and importantly what mechanisms act to organize the brain states themselves. Specifically, we lack a mechanistic explanation as to how task-independent shifts in neural activity arise (McGinley et al. 2015b).

We propose that a major source of trial-to-trial variability resides in a fundamental biological system, the arousal system, which is known to drive widespread and rapid changes in the dynamics of cortical networks (Coull 1998; Raut et al. 2021; Whyte 1992). We hypothesize that rapid changes in **arousal states** modulate broad cortical activity patterns (brain states) independent of task demands and stimulus characteristics, which can lead to substantial variability in perception and behavior. Crucially, the impact of arousal states occurs in concert with other sources of variability, including top-down processes like

categorical perception and task-dependent processes like attention. The rest of this review focuses on the putative role of arousal states in behavioral and perceptual variability.

5. Arousal States Modulate Brain States That Influence Perception

While it is clear that arousal can have important effects on neural activity, its effects on speech perception are not clearly understood. In particular, we lack a fundamental understanding of the links between arousal states (brainstem/LC-NE), cortical brain states (distributed network activity), and trial-by-trial variability in perceptual behavior (speech cortex/STG). We propose that arousal states modulate perceptual behavior and performance during auditory processing tasks by modulating brain states that influence perception (Figure 2).

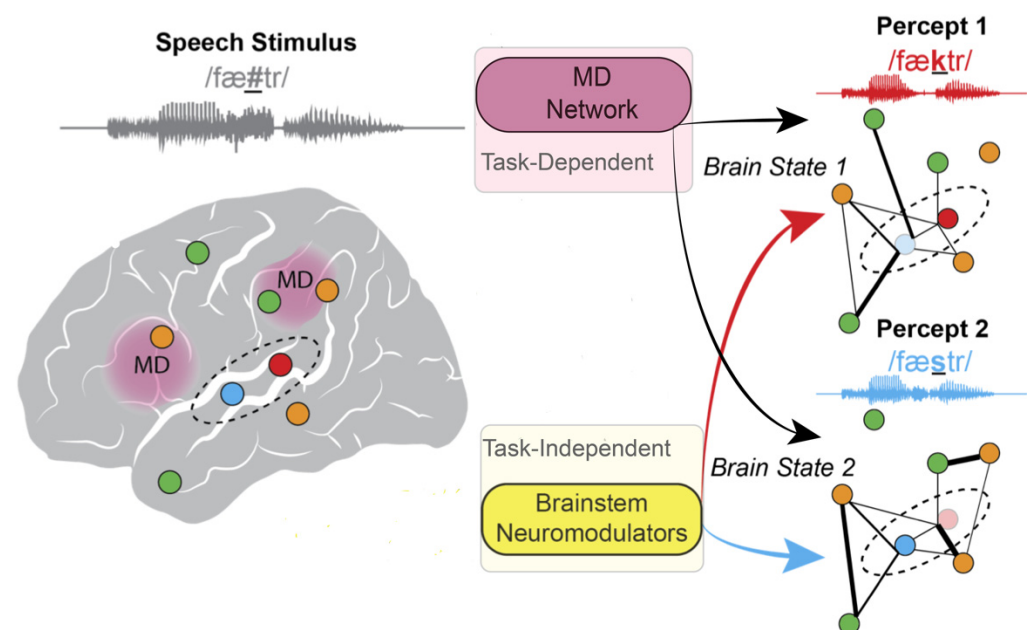


Figure 2. Trial-by-trial variability in perception is modulated by cortical brain states driven by arousal. Arousal state is mediated by the release of brainstem neuromodulators (such as norepinephrine; yellow box). Changes in arousal alter activity in numerous cortical sites including core speech cortex/STG (blue and red circles), non-core speech processing regions (green circles), and cortical structures comprising the Multiple Demand (MD) network (orange circles). An ambiguous stimulus (e.g., /fæ#tə/) may be perceived in multiple ways depending on moment-to-moment variability in brain states, the strength and configuration of functional networks (denoted by line thickness), and activity within cortical nodes tuned to specific features (shaded blue and red circles).

In human experiments, fluctuations in indices commonly associated with phasic arousal have been found to correlate with variability in perceptual outcomes, including the phase of scalp-recorded alpha oscillations. With regards to visual perception, pre-stimulus alpha phase has been shown to predict whether or not phosphenes (visual sensations perceived in the absence of an evoking stimulus) will be induced by transcranial magnetic stimulation (Dugue et al. 2011). Likewise, Strauss et al. (2015) found that correct and incorrect responses for a lexical decision task in noise could be predicted from prestimulus alpha phase. The authors suggest that in correct responses, the onset of the initial phoneme of the stimulus coincided with an optimal excitatory phase in which a perceptual object can be ‘selected’. Whether this effect is driven by top-down factors like attention or by subcortical arousal mechanisms remains unclear.

As stated in Section 2, levels of arousal are often separated into ‘low’, ‘moderate’, and ‘high’ states. Optimal performance in auditory tasks appears to coincide with moderate arousal states (McGinley et al. 2015a). For example, in an auditory detection task, listeners

attempted to detect the presence of a faint pure tone embedded in noise (de Gee et al. 2020). When tones were presented on 50% of trials, participants exhibited increased sensitivity and decreased reaction time when in a higher state of arousal. Similarly, optimal sensitivity to pitch differences in an auditory judgment task has been found to correspond to moderate levels of arousal (Waschke et al. 2019).

These findings indicate two potential mechanisms by which arousal may modulate brain states to affect perception and behavior. First, moderate, ‘optimal’ levels of arousal may decrease the spontaneous firing rate of specific populations to a greater extent than evoked activity (i.e., increase the activity within specific network nodes), thereby increasing the signal-to-noise ratio (SNR) of a signal (McBurney-Lin et al. 2019). Second, arousal may modulate the functional connectivity of broad cortical networks. Traveling waves of activity have been linked to fluctuations in arousal (Raut et al. 2021) and different levels of arousal have been found to correlate with changes in within- and between- network connectivity (Young et al. 2017). Given the distributed interactional nature of the processes involved in speech processing and language learning (Evans and Davis 2015; Feng et al. 2021; Friederici 2012) it is likely that these processes are affected by arousal-driven fluctuations in global brain states.

LC-NE mediated arousal has also been shown to be a key driver of neural plasticity, a key concept in learning (Marzo et al. 2009). The LC becomes active in response to a novel/salient stimulus, though this novelty response quickly habituates. However, if reinforcement is provided (e.g., in the form of a task reward) then the LC response re-emerges (Sara et al. 1994).

Given these robust effects on auditory processing and plasticity, we propose that changes in arousal state can explain variability in neural activity corresponding to performance in speech perception and learning. While not previously considered through the lens of arousal states, the mechanisms discussed here potentially explain the trial-by-trial variability observed in complex second language acquisition tasks like speech sound category learning (Yi et al. 2021). For example, it may be the case that differences in cortical responses to specific tones on correct versus incorrect trials reflect differences in SNR driven by rapid fluctuations in arousal (McBurney-Lin et al. 2019), with more accurate responses coinciding with moderate levels of arousal (Aston-Jones and Cohen 2005; de Gee et al. 2020). Alternatively, specific neural populations, primarily in bilateral STG and ventrolateral frontal cortex but also in broadly distributed regions within the MD network, exhibit shifts in neural response profiles in correspondence with learning. These shifts may reflect the influence of arousal state on cortical networks supporting tone categorization (Feng et al. 2021; Sara and Bouret 2012). Likewise, phasic arousal mechanisms may explain some of the stochastic behavior observed in multistable perception tasks (Leonard et al. 2016) and may also provide the modulatory effects necessary to allow top-down cognitive factors to influence perception (Holdgraf et al. 2016).

6. Using Non-Invasive Vagus Nerve Stimulation to Study the Effects of Arousal in Speech Perception and Learning

Thus far, evidence indicates that apparently random fluctuations in behavior and perception may be partially explainable by changes in arousal states. Perhaps the strongest evidence for the role of subcortically-mediated arousal states in perception and learning comes from studies that use causal manipulations of these putative mechanisms. In this section, we discuss a key method that will enable us to develop a mechanistic understanding of how arousal states influence speech sound perception and sound learning: transcutaneous auricular vagus nerve stimulation (taVNS).

Recent work has identified a relatively simple method for modulating arousal states via electrical stimulation. The peripheral nerves, and in particular the cranial nerves that project directly to the brainstem, can be targeted for modulating central nervous system activity (Adair et al. 2020; Bari and Pouratian 2012; Ginn et al. 2019). The vagus nerve in particular has been shown to be a key peripheral nerve associated with arousal and

cognition, in part due to its widespread connectivity to a variety of systems throughout the brain and body (Vonck et al. 2014; Vonck and Larsen 2018). Traditionally, vagus nerve stimulation (VNS) involves the implantation of a cuff electrode around the cervical vagus nerve in the neck and a signal generator/battery in the chest (iVNS). More recently, a non-surgical alternative has been developed targeting the auricular branch of the vagus nerve using transcutaneous surface electrodes (taVNS; Frangos et al. 2015; Ventureyra 2000; Figure 3a,b).

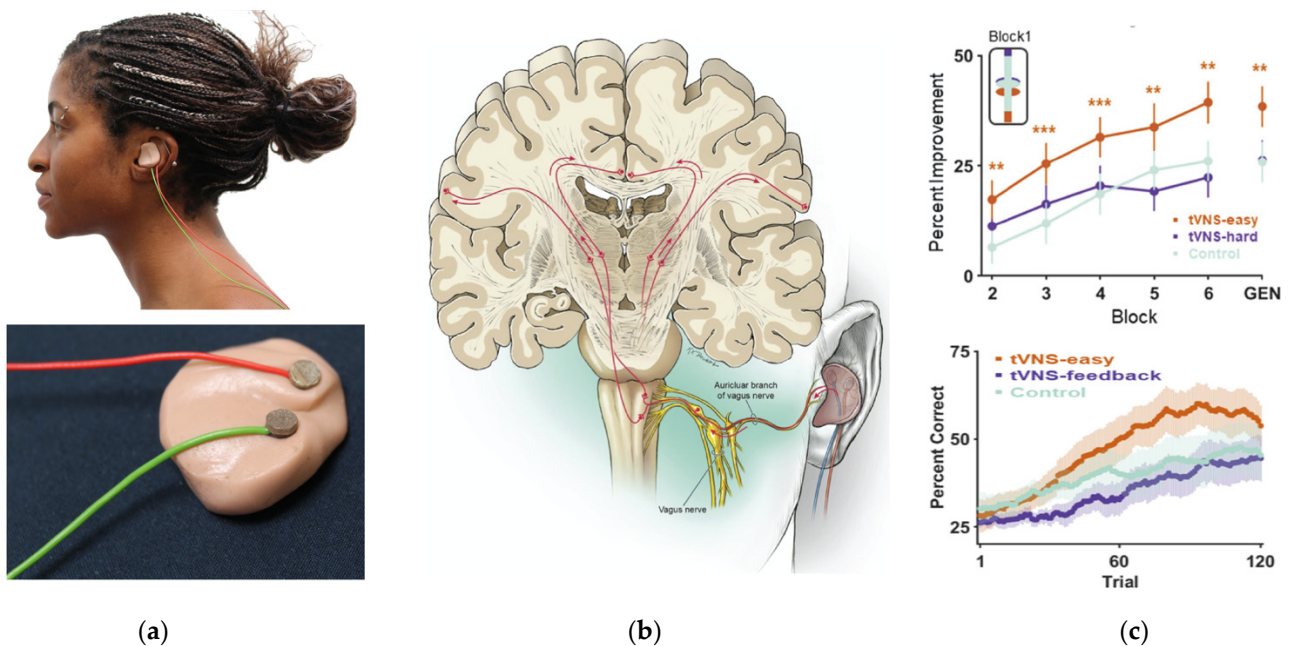


Figure 3. Enhancement of non-native speech sound learning using transcutaneous auricular vagus nerve stimulation (taVNS). (a) Electrical stimulation of the auricular branch of the vagus nerve (ABVN) using simple 4 mm Silver/Silver-Chloride embedded in silicone putty. Source: Image credited to Leonard Lab/UCSF/Jhia Louise Nicole Jackson; labeled for reuse. (b) Stimulation of the ABVN activates subcortical structures, such as the locus coeruleus (LC), which modulate the widespread release of neurotransmitters like norepinephrine (NE) throughout the cortex, including speech cortex and the MD network. (c) In a Mandarin tone learning task (Llanos et al. 2020), taVNS delivered during the presentation of easy-to-learn stimuli enhanced performance compared to control (top panel; orange and light blue lines). However, no effect was found when stimulation was delivered during feedback (bottom panel; purple line). ** and *** indicate increasing levels of statistical significance in a linear mixed effects model.

Application of taVNS involves delivering electrical activity to branches of the vagus nerves innervating the skin of the outer ear. Typically, electrodes are affixed to the target site using either an earbud (e.g., Frangos et al. 2015), clip (e.g., Fang et al. 2016), or moldable putty (e.g., Llanos et al. 2020; Schuerman et al. 2021). Stimulation patterns (i.e., the shape of the electrical waveform) can vary widely with regard to pulse shape (e.g., mono/biphasic square wave pulse), pulse amplitude, pulse width, and frequency (i.e., pulse delivery rate; for detailed reviews, see Kaniusas et al. 2019a; Yap et al. 2020). As evidenced in both animal (Hulsey et al. 2017; Loerwald et al. 2018; Morrison et al. 2020; Van Lysebettens et al. 2020) and human physiology (Badran et al. 2018; Schuerman et al. 2021; Yakunina et al. 2017), varying stimulation parameters can produce different effects, which will likely necessitate matching specific stimulation parameters to target outcomes.

While there are likely differences in innervation between the auricular and cervical pathways (Butt et al. 2019; Cakmak 2019), exploratory studies have revealed that taVNS may be able to achieve comparable effects as iVNS without the need for surgery (Kaniusas et al. 2019b; Schuerman et al. 2021). Crucially, both iVNS and taVNS have recently been

shown to modulate complex perceptual, motor, and cognitive processes, such as tonotopy (Engineer et al. 2011; Shetake et al. 2012), somatotopy (Darrow et al. 2020; Pruitt et al. 2016), auditory stimulus-reward association (Lai and David 2021), and responses to speech sounds (Engineer et al. 2015).

These effects have recently been extended into the domain of non-native sound category learning. In a recent study, participants were trained to recognize Mandarin tones while receiving taVNS (Llanos et al. 2020). Participants were randomly assigned to three stimulation groups: the first received taVNS on two ‘easy-to-learn’ tones that could be differentiated on the basis of pitch height (tones 1 and 3; ‘taVNS-easy’); the second received taVNS on two ‘hard-to-learn’ tones that differed by the direction of pitch change (tones 2 and 4; ‘taVNS-hard’); the third received no stimulation (‘Control’). During training, participants in the stimulation groups received peri-stimulus taVNS aligned to the onset of the auditory stimulus. Accuracy was found to be greater in the taVNS-easy group compared to both the taVNS-hard group and the control group, as well as compared to a normative aggregate sample of 678 comparable listeners (Figure 3c, top). Furthermore, increases in accuracy were specific to the stimulated tones. This study demonstrates that it is possible to rapidly modulate a key component of L2 learning using taVNS.

This study also revealed that, along with choice of stimulation parameters, the timing of stimulation relative to task events is likely to be a key consideration for the implementation and optimization of taVNS. Llanos et al. (2020) employed a peri-stimulus paradigm in which taVNS was aligned to the onset of the auditory stimulus. However, in a follow-up experiment, no learning enhancement was found when stimulation was paired with feedback on each trial (Figure 3c, bottom). Two recent studies on Mandarin word learning that directly compared peri-stimulus to continuously delivered taVNS found that both forms of taVNS improved performance. However, the patterns of effects differed between paradigms. For example, with regards to behavior, peri-stimulus taVNS was associated with increased accuracy on mismatch trials, whereas continuous stimulation was associated with a greater reduction in reaction time (Pandža et al. 2020; Phillips et al. 2021). Changes in tone-evoked pupillary responses on subsequent training days were greater for participants in the peri-stim group compared to sham or continuous stimulation (Pandža et al. 2020). Interestingly, the relationship between accuracy and the amplitude of the N400 EEG event-related potential was found to differ between peri-stim and continuous stimulation, with sham and continuous stimulation patterning together while peri-stimulus exhibited an inverse relationship from the two (Phillips et al. 2021). Further research in this area is required to determine whether these differences in behavior and physiology stem from dissociable mechanisms.

It remains to be established whether these types of learning enhancement effects are directly related to regulation of arousal states. The neuromodulatory pathway of VNS is believed to overlap extensively with subcortical structures regulating arousal. The vagus nerve innervates the LC-NE system, as well as several others, via projections to the nucleus of the solitary tract (Van Bockstaele et al. 1999; Ruggiero et al. 2000). Accordingly, activity in the LC has been found to increase in response to VNS stimulation (Hulsey et al. 2017), even at stimulation amplitudes as low as 0.1 milliamperes. Activation of the LC-NE system by VNS is also supported by research in non-human animals that has found rapid, dose-dependent increases in pupil dilation and cortical activation in response to VNS (Collins et al. 2021; Mridha et al. 2021) and similar findings have been reported in humans using i/taVNS (Desbeaumes Jodoin et al. 2015; Pandža et al. 2020; Sharon et al. 2021; Urbin et al. 2021; though c.f. Burger et al. 2020; D’Agostini et al. 2021; Schevernels et al. 2016; Warren et al. 2019). taVNS has also been found to modulate other biomarkers of LC-NE activity and arousal, such as alpha oscillations (Sharon et al. 2021) and salivary alpha-amylase and cortisol levels (Warren et al. 2019). Finally, VNS has been found to elicit activation across distributed cortical regions, consistent with widespread neurotransmitter release (Cao et al. 2017; Schuerman et al. 2021). These findings suggest that rapid effects of VNS on performance are likely mediated by the arousal system.

Given the established links between VNS and arousal, we propose that the modulation of performance observed in this and other studies reflects the targeted modulation of arousal states during the learning process. When taVNS is paired with a specific task or behavior that engages a particular set of brain regions (e.g., STG during a speech perception task), changes in domain-general arousal can lead to reinforced activity patterns in core task-relevant areas (Engineer et al. 2019; Hulsey et al. 2016). The widespread release of arousal-modulating neurotransmitters may allow representations that are stored in cortical circuits to be either enhanced or disrupted, depending on factors like timing and behavioral task parameters (Berridge and Waterhouse 2003). For example, the identity of the phoneme perceived in a perceptual restoration task can be predicted by activity in broader cortical networks, including inferior frontal cortex, up to ~300 milliseconds before the onset of the sound (Leonard et al. 2016), suggesting that activity in broadly-distributed areas influences perception in a top-down fashion. Similarly, improvement in Mandarin tone learning has been found to correspond to dynamic changes in activity across the MD network (Feng et al. 2021). Rapid changes in arousal driven by VNS may act to alter the configuration of such networks during speech perception, suppressing or enhancing the activity of neural populations with response properties relevant to tone learning (Yi et al. 2021). More specifically, taVNS-induced arousal may modulate the SNR of task-relevant neural populations (i.e., those tuned to specific features, such as changes in pitch; McBurney-Lin et al. 2019). While there is more work necessary to establish these links directly, we propose that this framework provides a set of testable hypotheses that may allow for these neuromodulatory systems to be incorporated into a new model of speech perception and learning that accounts for within- as well as across- individual variability (Figure 2).

To establish this link more explicitly requires reliable biomarkers that reflect changes in arousal state as well as methods for modulating the activity of the arousal system in a causal manner. It has been known for more than half a century that changes in pupil diameter reflect cognitive load and affective arousal (Kahneman and Beatty 1966; Stanners et al. 1979). More recent research has revealed rapid fluctuations in pupillary responses depending on stimulus properties, behavior, and task demands (de Gee et al. 2017, 2020; Gilzenrat et al. 2010; McGinley et al. 2015a; Reimer et al. 2016). While pupil responses appear to be influenced by multiple subcortical pathways (Berridge and Waterhouse 2003; Larsen and Waters 2018), changes in pupil diameter show robust correlations with activity in the LC (Joshi et al. 2016). Furthermore, both electrical (Liu et al. 2017) and optogenetic (Breton-Provencher and Sur 2019) stimulation of the LC increases pupil diameter. These findings suggest that activity in the LC-NE system influences, if not drives, rapid fluctuations in pupil diameter (Joshi et al. 2016). Overall, the evidence indicates that pupil dilation is a strong biomarker for LC-NE mediated changes in arousal state.

7. Conclusions

In this review, we have discussed the numerous factors that influence trial-by-trial variability in speech perception and learning tasks. Studies that employ invasive electrophysiological techniques or combine non-invasive imaging with advanced computational methods are rapidly generating new insights into the neural foundations of perception and non-native speech sound learning. Together, these recent advances have begun to explain the sources of neural variability underlying behavioral differences between individuals. Furthermore, these techniques make it possible to investigate the timecourse of perception and learning within a single individual, generating novel insights regarding how the dynamics of neural activity reflect stimulus- (e.g., context) or task- (e.g., attention) related variability.

However, it is clear that there is a substantial portion of this variability that has not been explained by traditionally studied factors like top-down cognitive modulation. Classic examples of multistable perception such as phoneme restoration clearly demonstrate that variability exists that is both stimulus- and task- independent. Given the challenges

listeners and learners face in real-world scenarios, understanding as much of this variability as possible may be crucial to developing an accurate model of perception, and for creating translational and pedagogical tools for improving outcomes in fields such as second language acquisition.

We propose that a more complete model of speech perception and learning should consider the role of subcortically-mediated arousal states. Emerging research is demonstrating how rapid fluctuations in arousal state can affect perceptual outcomes as well as related behavior. To determine how arousal states fit into a more complete model of perception, it is crucial that we not only are able to track correlations between arousal and perception, but also able to manipulate arousal states in order to identify causal links between the two. In this regard, taVNS constitutes a novel, promising tool for studying the brain systems that underlie these mechanisms. With such tools, it is increasingly feasible to integrate long-overlooked systems into our thinking about complex behaviors like speech perception and non-native sound learning. We are excited for the next several years of research on this topic and are optimistic that this work will contribute to major advances in our understanding.

Author Contributions: Conceptualization, W.L.S., B.C. and M.K.L. writing—original draft preparation, W.L.S., B.C. and M.K.L.; writing—review and editing, W.L.S. and M.K.L.; visualization, W.L.S. and M.K.L.; supervision, M.K.L.; funding acquisition, W.L.S., B.C. and M.K.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Defense Advanced Research Projects Agency contract no. N66001-17-2-4008, NIH R01-DC0155004, and NWO Rubicon Grant 446-17-002.

Institutional Review Board Statement: Not Applicable.

Informed Consent Statement: Not Applicable.

Data Availability Statement: Not Applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Adair, Devin, Dennis Truong, Zeinab Esmaeilpour, Nigel Gebodh, Helen Borges, Libby Ho, and J. Douglas Bremner. 2020. Electrical stimulation of cranial nerves in cognition and disease. *Brain Stimulation* 13: 717–50. [\[CrossRef\]](#) [\[PubMed\]](#)
- Aston-Jones, Gary, and Jonathan D. Cohen. 2005. An integrative theory of locus coeruleus-norepinephrine function: Adaptive Gain and Optimal Performance. *Annual Review of Neuroscience* 28: 403–50. [\[CrossRef\]](#)
- Badran, Bashar W., Oliver J. Mithoefer, Caroline E. Summer, Nicholas T. LaBate, Chloe E. Glusman, Alan W. Badran, and William H. DeVries. 2018. Short trains of transcutaneous auricular vagus nerve stimulation (taVNS) have parameter-specific effects on heart rate. *Brain Stimulation* 11: 699–708. [\[CrossRef\]](#) [\[PubMed\]](#)
- Bari, Ausaf A., and Nader Pouratian. 2012. Brain imaging correlates of peripheral nerve stimulation. *Surgical Neurology International* 3: 260. [\[CrossRef\]](#)
- Berridge, Craig W., and Barry D. Waterhouse. 2003. The locus coeruleus–noradrenergic system: Modulation of behavioral state and state-dependent cognitive processes. *Brain Research Reviews* 42: 33–84. [\[CrossRef\]](#)
- Birdsong, David. 2018. Plasticity, variability and age in second language acquisition and bilingualism. *Frontiers in Psychology* 9: 81. [\[CrossRef\]](#) [\[PubMed\]](#)
- Birdsong, David, and Jan Vanhove. 2016. Age of second language acquisition: Critical periods and social concerns. In *Bilingualism across the Lifespan: Factors Moderating Language Proficiency*. Language and the Human Lifespan Series; Washington, DC: American Psychological Association, pp. 163–81. [\[CrossRef\]](#)
- Bradlow, Ann R. 2008. Training non-native language sound patterns: Lessons from training Japanese adults on the English /r/-/l/ contrast. In *Phonology and Second Language Acquisition*. Edited by Jette G. Hansen Edwards and Mary L. Zampini. Amsterdam: John Benjamins Publishing Company, pp. 287–308.
- Bradlow, Ann R., and Tessa Bent. 2008. Perceptual adaptation to non-native speech. *Cognition* 106: 707–29. [\[CrossRef\]](#)
- Breton-Provencher, Vincent, and Mriganka Sur. 2019. Active control of arousal by a locus coeruleus GABAergic circuit. *Nature Neuroscience* 22: 218–28. [\[CrossRef\]](#)
- Brodbeck, Christian, Alex Jiao, L. Elliot Hong, and Jonathan Z. Simon. 2020. Neural speech restoration at the cocktail party: Auditory cortex recovers masked speech of both attended and ignored speakers. *PLOS Biology* 18: e3000883. [\[CrossRef\]](#) [\[PubMed\]](#)
- Brosseau-Lapr , Fran oise, Susan Rvachew, Meghan Clayards, and Daniel Dickson. 2013. Stimulus variability and perceptual learning of nonnative vowel categories. *Applied Psycholinguistics* 34: 419–41. [\[CrossRef\]](#)

- Brouwer, Susanne, Holger Mitterer, and Falk Huettig. 2013. Discourse context and the recognition of reduced and canonical spoken words. *Applied Psycholinguistics* 34: 519–39. [\[CrossRef\]](#)
- Burger, Andreas Michael, Willem A. J. van der Does, Jos F. Brosschot, and Bart Verkuil. 2020. From ear to eye? No effect of transcutaneous vagus nerve stimulation on human pupil dilation: A report of three studies. *Biological Psychology* 152: 107863. [\[CrossRef\]](#)
- Butt, Mohsin F., Ahmed Albusoda, Adam D. Farmer, and Qasim Aziz. 2019. The anatomical basis for transcutaneous auricular vagus nerve stimulation. *Journal of Anatomy* 236: 588–611. [\[CrossRef\]](#)
- Cakmak, Yusuf Ozgur. 2019. Concerning auricular vagal nerve stimulation: Occult neural networks. *Frontiers in Human Neuroscience* 13: 421. [\[CrossRef\]](#)
- Campbell, Ruth. 2008. The processing of audio-visual speech: Empirical and neural bases. *Philosophical Transactions of the Royal Society B: Biological Sciences* 363: 1001–10. [\[CrossRef\]](#) [\[PubMed\]](#)
- Cao, Jiayue, Kun-Han Lu, Terry L. Powley, and Zhongming Liu. 2017. Vagal nerve stimulation triggers widespread responses and alters large-scale functional connectivity in the rat brain. *PLoS ONE* 12: e0189518. [\[CrossRef\]](#)
- Cardin, Jessica A., and Marc F. Schmidt. 2003. Song system auditory responses are stable and highly tuned during sedation, rapidly modulated and unselective during wakefulness, and suppressed by arousal. *Journal of Neurophysiology* 90: 2884–99. [\[CrossRef\]](#) [\[PubMed\]](#)
- Cardin, Jessica A., and Marc F. Schmidt. 2004. Noradrenergic inputs mediate state dependence of auditory responses in the avian song system. *Journal of Neuroscience* 24: 7745–53. [\[CrossRef\]](#)
- Chandrasekaran, Bharath, Han-Gyol Yi, Kirsten E. Smayda, and W. Todd Maddox. 2016. Effect of explicit dimensional instruction on speech category learning. *Attention, Perception, & Psychophysics* 78: 566–82. [\[CrossRef\]](#)
- Chandrasekaran, Bharath, Padma D. Sampath, and Patrick C. M. Wong. 2010. Individual variability in cue-weighting and lexical tone learning. *The Journal of the Acoustical Society of America* 128: 456–65. [\[CrossRef\]](#)
- Chandrasekaran, Bharath, Seth R. Koslov, and W. Todd Maddox. 2014. Toward a dual-learning systems model of speech category learning. *Frontiers in Psychology* 5: 825. [\[CrossRef\]](#) [\[PubMed\]](#)
- Chang, Edward F. 2015. Towards large-scale, human-based, mesoscopic neurotechnologies. *Neuron* 86: 68–78. [\[CrossRef\]](#)
- Collins, Lindsay, Laura Boddington, Paul J. Steffan, and David McCormick. 2021. Vagus nerve stimulation induces widespread cortical and behavioral activation. *Current Biology* 31: 2088–98. [\[CrossRef\]](#)
- Coull, Jennifer T. 1998. Neural correlates of attention and arousal: Insights from electrophysiology, functional neuroimaging and psychopharmacology. *Progress in Neurobiology* 55: 343–61. [\[CrossRef\]](#)
- Coull, Jennifer T., Christian Büchel, Karl J. Friston, and Chris D. Frith. 1999. Noradrenergically mediated plasticity in a human attentional neuronal network. *NeuroImage* 10: 705–15. [\[CrossRef\]](#) [\[PubMed\]](#)
- D'Agostini, Martina, Andreas M. Burger, Mathijs Franssen, Nathalie Claes, Mathias Weymar, Andreas von Leupoldt, and Ilse Van Diest. 2021. Effects of transcutaneous auricular vagus nerve stimulation on reversal learning, tonic pupil size, salivary alpha-amylase, and cortisol. *Psychophysiology* 58: e13885. [\[CrossRef\]](#)
- Darrow, Michael J., Miranda Torres, Maria J. Sosa, Tanya T. Danaphongse, Zainab Haider, Robert L. Rennaker, Michael P. Kilgard, and Seth A. Hays. 2020. Vagus nerve stimulation paired with rehabilitative training enhances motor recovery after bilateral spinal cord injury to cervical forelimb motor pools. *Neurorehabilitation and Neural Repair* 34: 200–209. [\[CrossRef\]](#) [\[PubMed\]](#)
- Davis, Matthew H., and Ingrid S. Johnsrude. 2007. Hearing speech sounds: Top-down influences on the interface between audition and speech perception. *Hearing Research* 229: 132–47. [\[CrossRef\]](#) [\[PubMed\]](#)
- Davis, Matthew H., Ingrid S. Johnsrude, Alexis Hervais-Adelman, Karen Taylor, and Carolyn McGettigan. 2005. Lexical information drives perceptual learning of distorted speech: Evidence from the comprehension of noise-vocoded sentences. *Journal of Experimental Psychology: General* 134: 222–41. [\[CrossRef\]](#)
- de Gee, Jan Willem, Konstantinos Tsetsos, Lars Schwabe, Anne E. Urai, David McCormick, Matthew J. McGinley, and Tobias H. Donner. 2020. Pupil-linked phasic arousal predicts a reduction of choice bias across species and decision domains. *eLife* 9: e54014. [\[CrossRef\]](#)
- de Gee, Jan Willem, Olympia Colizoli, Niels A. Kloosterman, Tomas Knapen, Sander Nieuwenhuis, and Tobias H. Donner. 2017. Dynamic modulation of decision biases by brainstem arousal systems. Edited by Klaas Enno Stephan. *eLife* 6: e23232. [\[CrossRef\]](#) [\[PubMed\]](#)
- Desbeaumes Jodoin, Véronique, Paul Lespérance, Dang K. Nguyen, Marie-Pierre Fournier-Gosselin, and Francois Richer. 2015. Effects of vagus nerve stimulation on pupillary function. *International Journal of Psychophysiology* 98: 455–59. [\[CrossRef\]](#) [\[PubMed\]](#)
- Diachek, Evgeniia, Idan Blank, Matthew Siegelman, Josef Affourtit, and Evelina Fedorenko. 2020. The domain-general Multiple Demand (MD) network does not support core aspects of language comprehension: A large-scale fMRI investigation. *Journal of Neuroscience* 40: 4536–50. [\[CrossRef\]](#) [\[PubMed\]](#)
- Ding, Nai, and Jonathan Z. Simon. 2012. Emergence of neural encoding of auditory objects while listening to competing speakers. *Proceedings of the National Academy of Sciences of USA* 109: 11854–59. [\[CrossRef\]](#)
- Dugue, Laura, Philippe Marque, and Rufin VanRullen. 2011. The phase of ongoing oscillations mediates the causal relation between brain excitation and visual perception. *Journal of Neuroscience* 31: 11889–93. [\[CrossRef\]](#)
- Duncan, John. 2010. The multiple-demand (MD) system of the primate brain: Mental programs for intelligent behaviour. *Trends in Cognitive Sciences* 14: 172–79. [\[CrossRef\]](#)

- Engineer, Navzer D., Jonathan R. Riley, Jonathan D. Seale, Will A. Vrana, Jai A. Shetake, Sindhu P. Sudanagunta, Michael S. Borland, and Michael P. Kilgard. 2011. Reversing pathological neural activity using targeted plasticity. *Nature* 470: 101–14. [\[CrossRef\]](#)
- Engineer, Crystal T., Navzer D. Engineer, Jonathan R. Riley, Jonathan D. Seale, and Michael P. Kilgard. 2015. Pairing speech sounds with vagus nerve stimulation drives stimulus-specific cortical plasticity. *Brain Stimulation* 8: 637–44. [\[CrossRef\]](#)
- Engineer, Navzer D., Teresa J. Kimberley, Cecilia N. Prudente, Jesse Dawson, W. Brent Tarver, and Seth A. Hays. 2019. Targeted vagus nerve stimulation for rehabilitation after stroke. *Frontiers in Neuroscience* 13: 280. [\[CrossRef\]](#)
- Erb, Julia, Molly J. Henry, Frank Eisner, and Jonas Obleser. 2013. The brain dynamics of rapid perceptual adaptation to adverse listening conditions. *Journal of Neuroscience* 33: 10688–97. [\[CrossRef\]](#)
- Evans, Samuel, and Matthew H. Davis. 2015. Hierarchical organization of auditory and motor representations in speech perception: Evidence from searchlight similarity analysis. *Cerebral Cortex* 25: 4772–88. [\[CrossRef\]](#) [\[PubMed\]](#)
- Faisal, A. Aldo, Luc P. J. Selen, and Daniel M. Wolpert. 2008. Noise in the nervous system. *Nature Reviews. Neuroscience* 9: 292–303. [\[CrossRef\]](#) [\[PubMed\]](#)
- Fang, Jiliang, Peijing Rong, Yang Hong, Yangyang Fan, Jun Liu, Honghong Wang, and Guolei Zhang. 2016. Transcutaneous vagus nerve stimulation modulates default mode network in major depressive disorder. *Biological Psychiatry* 79: 266–73. [\[CrossRef\]](#)
- Feng, Gangyi, Han Gyol Yi, and Bharath Chandrasekaran. 2019. The role of the human auditory corticostriatal network in speech learning. *Cerebral Cortex* 29: 4077–89. [\[CrossRef\]](#)
- Feng, Gangyi, Zhenzhong Gan, Fernando Llanos, Danting Meng, Suiping Wang, Patrick C. M. Wong, and Bharath Chandrasekaran. 2021. A distributed dynamic brain network mediates linguistic tone representation and categorization. *NeuroImage* 224: 117410. [\[CrossRef\]](#)
- Finn, Amy Sue, Carla L. Hudson Kam, Marc Ettlinger, Jason Vytlaçil, and Mark D'Esposito. 2013. Learning language with the wrong neural scaffolding: The cost of neural commitment to sounds. *Frontiers in Systems Neuroscience* 7: 85. [\[CrossRef\]](#)
- Fox, Michael D., Abraham Z. Snyder, Justin L. Vincent, and Marcus E. Raichle. 2007. Intrinsic fluctuations within cortical systems account for intertrial variability in human behavior. *Neuron* 56: 171–84. [\[CrossRef\]](#)
- Frangos, Eleni, Jens Ellrich, and Barry R. Komisaruk. 2015. Non-invasive access to the vagus nerve central projections via electrical stimulation of the external ear: fMRI evidence in humans. *Brain Stimulation* 8: 624–36. [\[CrossRef\]](#) [\[PubMed\]](#)
- Friederici, Angela D. 2012. The cortical language circuit: From auditory perception to sentence comprehension. *Trends in Cognitive Sciences* 16: 262–68. [\[CrossRef\]](#)
- Ganong, William F. 1980. Phonetic categorization in auditory word perception. *Journal of Experimental Psychology. Human Perception and Performance* 6: 110–25. [\[CrossRef\]](#)
- Gilzenrat, Mark S., Sander Nieuwenhuis, Marieke Jepma, and Jonathan D. Cohen. 2010. Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive, Affective, & Behavioral Neuroscience* 10: 252–69. [\[CrossRef\]](#)
- Ginn, Claire, Bipin Patel, and Robert Walker. 2019. Existing and emerging applications for the neuromodulation of nerve activity through targeted delivery of electric stimuli. *International Journal of Neuroscience* 129: 1013–23. [\[CrossRef\]](#)
- Golestani, Narly, and Robert J. Zatorre. 2009. Individual differences in the acquisition of second language phonology. *Brain and Language* 109: 55–67. [\[CrossRef\]](#)
- Guediche, Sara, Sheila Blumstein, Julie Fiez, and Lori Holt. 2014. Speech perception under adverse conditions: Insights from behavioral, computational, and neuroscience research. *Frontiers in Systems Neuroscience* 7: 126. [\[CrossRef\]](#)
- Harris, Kenneth D., and Alexander Thiele. 2011. Cortical state and attention. *Nature Reviews Neuroscience* 12: 509–23. [\[CrossRef\]](#)
- Hasson, Uri, Giovanna Egidi, Marco Marelli, and Roel M. Willems. 2018. Grounding the neurobiology of language in first principles: The necessity of non-language-centric explanations for language comprehension. *Cognition* 180: 135–57. [\[CrossRef\]](#)
- Heald, Shannon, and Howard Charles Nusbaum. 2014. Speech perception as an active cognitive process. *Frontiers in Systems Neuroscience* 8: 1–15. [\[CrossRef\]](#)
- Holdgraf, Christopher R., Wendy de Heer, Brian Pasley, Jochem Rieger, Nathan Crone, Jack J. Lin, Robert T. Knight, and Frédéric E. Theunissen. 2016. Rapid tuning shifts in human auditory cortex enhance speech intelligibility. *Nature Communications* 7: 13654. [\[CrossRef\]](#) [\[PubMed\]](#)
- Hulsey, Daniel R., Jonathan R. Riley, Kristofer W. Loerwald, Robert L. Rennaker II, Michael P. Kilgard, and Seth A. Hays. 2017. Parametric characterization of neural activity in the locus coeruleus in response to vagus nerve stimulation. *Experimental Neurology* 289: 21–30. [\[CrossRef\]](#)
- Hulsey, Daniel R., Seth A. Hays, Navid Khodaparast, Andrea Ruiz, Priyanka Das, Robert L. Rennaker II, and Michael P. Kilgard. 2016. Reorganization of motor cortex by vagus nerve stimulation requires cholinergic innervation. *Brain Stimulation* 9: 174–81. [\[CrossRef\]](#) [\[PubMed\]](#)
- Huyck, Julia Jones, and Ingrid S. Johnsrude. 2012. Rapid perceptual learning of noise-vocoded speech requires attention. *The Journal of the Acoustical Society of America* 131: EL236–42. [\[CrossRef\]](#) [\[PubMed\]](#)
- Joshi, Siddhartha, Yin Li, Rishi Kalwani, and Joshua I. Gold. 2016. Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. *Neuron* 89: 221–34. [\[CrossRef\]](#) [\[PubMed\]](#)
- Kahneman, Daniel, and Jackson Beatty. 1966. Pupil diameter and load on memory. *Science* 154: 1583–85. [\[CrossRef\]](#)
- Kang, Shinae, Keith Johnson, and Gregory Finley. 2016. Effects of native language on compensation for coarticulation. *Speech Communication* 77: 84–100. [\[CrossRef\]](#)

- Kaniusas, Eugenijus, Stefan Kampusch, Marc Tittgemeyer, Fivos Panetsos, Raquel Fernandez Gines, Michele Papa, Attila Kiss, Bruno Podesser, Antonino Mario Cassara, Emmeric Tanghe, and et al. 2019a. Current directions in the auricular vagus nerve Stimulation I—A physiological perspective. *Frontiers in Neuroscience* 13: 1–23. [\[CrossRef\]](#)
- Kaniusas, Eugenijus, Stefan Kampusch, Marc Tittgemeyer, Fivos Panetsos, Raquel Fernandez Gines, Michele Papa, Attila Kiss, Bruno Podesser, Antonino Mario Cassara, Emmeric Tanghe, and et al. 2019b. Current directions in the auricular vagus nerve stimulation II—An engineering perspective. *Frontiers in Neuroscience* 13: 1–16. [\[CrossRef\]](#)
- Khoshkhou, Sattar, Matthew K. Leonard, Nima Mesgarani, and Edward F. Chang. 2018. Neural correlates of sine-wave speech intelligibility in human frontal and temporal cortex. *Brain and Language* 187: 83–91. [\[CrossRef\]](#)
- Kleinow, Jennifer, and Anne Smith. 2006. Potential interactions among linguistic, autonomic, and motor factors in speech. *Developmental Psychobiology* 48: 275–87. [\[CrossRef\]](#)
- Kral, Andrej, Michael F. Dorman, and Blake S. Wilson. 2019. Neuronal development of hearing and language: Cochlear implants and critical periods. *Annual Review of Neuroscience* 42: 47–65. [\[CrossRef\]](#) [\[PubMed\]](#)
- Kuhl, Patricia K. 2004. Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience* 5: 831–43. [\[CrossRef\]](#)
- Kuhl, Patricia K. 2010. Brain mechanisms in early language acquisition. *Neuron* 67: 713–27. [\[CrossRef\]](#) [\[PubMed\]](#)
- Kuhl, Patricia K., Barbara T. Conboy, Denise Padden, Tobey Nelson, and Jessica Pruitt. 2005. Early speech perception and later language development: Implications for the “critical period”. *Language Learning and Development* 1: 237–64. [\[CrossRef\]](#)
- Kutas, Marta, and Kara D. Federmeier. 2000. Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences* 4: 463–70. [\[CrossRef\]](#)
- Lai, Jesyin, and Stephen V. David. 2021. Short-term effects of vagus nerve stimulation on learning and evoked activity in auditory cortex. *ENeuro*, 8. [\[CrossRef\]](#)
- Larsen, Rylan S., and Jack Waters. 2018. Neuromodulatory correlates of pupil dilation. *Frontiers in Neural Circuits* 12: 21. [\[CrossRef\]](#) [\[PubMed\]](#)
- Leonard, Matthew K., Maxime O. Baud, Matthias J. Sjerps, and Edward F. Chang. 2016. Perceptual restoration of masked speech in human cortex. *Nature Communications* 7: 13619. [\[CrossRef\]](#)
- Lim, Sung-joo, and Lori L. Holt. 2011. Learning foreign sounds in an alien world: Videogame training improves non-native speech categorization. *Cognitive Science* 35: 1390–405. [\[CrossRef\]](#) [\[PubMed\]](#)
- Lin, Pei-Ann, Samuel K. Asinof, Nicholas J. Edwards, and Jeffry S. Isaacson. 2019. Arousal regulates frequency tuning in primary auditory cortex. *Proceedings of the National Academy of Sciences of USA* 116: 25304–10. [\[CrossRef\]](#)
- Liu, Yang, Charles Rodenkirch, Nicole Moskowitz, Brian Schriver, and Qi Wang. 2017. Dynamic lateralization of pupil dilation evoked by locus coeruleus activation results from sympathetic, not parasympathetic, contributions. *Cell Reports* 20: 3099–112. [\[CrossRef\]](#)
- Llanos, Fernando, Jacie R. McHaney, William L. Schuerman, Han G. Yi, Matthew K. Leonard, and Bharath Chandrasekaran. 2020. Non-invasive peripheral nerve stimulation selectively enhances speech category learning in adults. *NPJ Science of Learning* 5: 1–11. [\[CrossRef\]](#)
- Loerwald, Kristofer W., Elizabeth P. Buell, Michael S. Borland, Robert L. Rennaker, Seth A. Hays, and Michael P. Kilgard. 2018. Varying stimulation parameters to improve cortical plasticity generated by VNS-tone pairing. *Neuroscience* 388: 239–47. [\[CrossRef\]](#)
- Luthra, Sahil, Giovanni Peraza-Santiago, Keiana Beeson, David Saltzman, Anne Marie Crinnion, and James S. Magnuson. 2021. Robust lexically mediated compensation for coarticulation: Christmash time is here again. *Cognitive Science* 45: e12962. [\[CrossRef\]](#)
- Maddox, W. Todd, and Bharath Chandrasekaran. 2014. Tests of a dual-system model of speech category learning*. *Bilingualism: Language and Cognition* 17: 709–28. [\[CrossRef\]](#)
- Mai, Guangting, Tim Schoof, and Peter Howell. 2019. Modulation of phase-locked neural responses to speech during different arousal states is age-dependent. *NeuroImage* 189: 734–44. [\[CrossRef\]](#)
- Mann, Virginia A., and Bruno H. Repp. 1980. Influence of vocalic context on perception of the [f]-[s] distinction. *Perception & Psychophysics* 28: 213–28. [\[CrossRef\]](#)
- Martins, Ana Raquel O., and Robert C. Froemke. 2015. Coordinated forms of noradrenergic plasticity in the locus coeruleus and primary auditory cortex. *Nature Neuroscience* 18: 1483–92. [\[CrossRef\]](#)
- Marzo, Aude, Jing Bai, and Satoru Otani. 2009. Neuroplasticity regulation by noradrenaline in mammalian brain. *Current Neuropharmacology* 7: 286–95. [\[CrossRef\]](#)
- Mattys, Sven L., F. Seymour, Angela S. Attwood, and Marcus R. Munafò. 2013. Effects of acute anxiety induction on speech perception: Are anxious listeners distracted listeners? *Psychological Science* 24: 1606–608. [\[CrossRef\]](#)
- McBurney-Lin, Jim, Ju Lu, Yi Zuo, and Hongdian Yang. 2019. Locus coeruleus-norepinephrine modulation of sensory processing and perception: A focused review. *Neuroscience & Biobehavioral Reviews* 105: 190–99. [\[CrossRef\]](#)
- McCormick, David A., and Hans-Christian Pape. 1990. Noradrenergic and serotonergic modulation of a hyperpolarization-activated cation current in thalamic relay neurones. *The Journal of Physiology* 431: 319–42. [\[CrossRef\]](#)
- McGinley, Matthew J., Martin Vinck, Jacob Reimer, Renata Batista-Brito, Edward Zagha, Cathryn R. Cadwell, Andreas S. Tolias, Jessica A. Cardin, and David A. McCormick. 2015a. Waking state: Rapid variations modulate neural and behavioral responses. *Neuron* 87: 1143–61. [\[CrossRef\]](#)
- McGinley, Matthew J., Stephen V. David, and David A. McCormick. 2015b. Cortical membrane potential signature of optimal states for sensory signal detection. *Neuron* 87: 179–92. [\[CrossRef\]](#)
- McGurk, Harry, and John MacDonald. 1976. Hearing lips and seeing voices. *Nature* 264: 746–48. [\[CrossRef\]](#)

- Menon, Vinod, and Sonia Crottaz-Herbette. 2005. Combined EEG and fMRI Studies of human brain function. In *International Review of Neurobiology*. Neuroimaging, Part A. Cambridge: Academic Press, vol. 66, pp. 291–321. [\[CrossRef\]](#)
- Mesgarani, Nima, and Edward F. Chang. 2012. Selective cortical representation of attended speaker in multi-talker speech perception. *Nature* 485: 233–36. [\[CrossRef\]](#)
- Mesgarani, Nima, Connie Cheung, Keith Johnson, and Edward F. Chang. 2014. Phonetic feature encoding in human superior temporal gyrus. *Science* 343: 1006–10. [\[CrossRef\]](#) [\[PubMed\]](#)
- Miller, George A., and Stephen Isard. 1963. Some perceptual consequences of linguistic rules. *Journal of Verbal Learning and Verbal Behavior* 2: 217–28. [\[CrossRef\]](#)
- Morrison, Robert A., Tanya T. Danaphongse, David T. Pruitt, Katherine S. Adcock, Jobin K. Mathew, Stephanie T. Abe, Dina M. Abdulla, Robert L. Rennaker, Michael P. Kilgard, and Seth A. Hays. 2020. A limited range of vagus nerve stimulation intensities produce motor cortex reorganization when delivered during training. *Behavioural Brain Research* 391: 112705. [\[CrossRef\]](#)
- Mridha, Zakir, Jan Willem de Gee, Yanchen Shi, Rayan Alkashgari, Justin Williams, Aaron Suminski, Matthew P. Ward, Wenhao Zhang, and Matthew James McGinley. 2021. Graded recruitment of pupil-linked neuromodulation by parametric stimulation of the vagus nerve. *Nature Communications* 12: 1539. [\[CrossRef\]](#) [\[PubMed\]](#)
- Myers, Emily B. 2014. Emergence of category-level sensitivities in non-native speech sound learning. *Frontiers in Neuroscience* 8: 1–11. [\[CrossRef\]](#) [\[PubMed\]](#)
- Norris, Dennis, James M. McQueen, and Anne Cutler. 2003. Perceptual learning in speech. *Cognitive Psychology* 47: 204–38. [\[CrossRef\]](#)
- Pandža, Nick B., Ian Phillips, Valerie P. Karuzis, Polly O'Rourke, and Stefanie E. Kuchinsky. 2020. Neurostimulation and pupillometry: New directions for learning and research in applied linguistics. *Annual Review of Applied Linguistics* 40: 56–77. [\[CrossRef\]](#)
- Paulon, Giorgio, Fernando Llanos, Bharath Chandrasekaran, and Abhra Sarkar. 2020. Bayesian semiparametric longitudinal drift-diffusion mixed models for tone learning in adults. *Journal of the American Statistical Association* 116: 1114–27. [\[CrossRef\]](#)
- Perrachione, Tyler K., Jiyeon Lee, Louise Y. Y. Ha, and Patrick C. M. Wong. 2011. Learning a novel phonological contrast depends on interactions between individual differences and training paradigm design. *The Journal of the Acoustical Society of America* 130: 461–72. [\[CrossRef\]](#)
- Phillips, Ian, Regina C. Calloway, Valerie P. Karuzis, Nick B. Pandža, Polly O'Rourke, and Stefanie E. Kuchinsky. 2021. Transcutaneous auricular vagus nerve stimulation strengthens semantic representations of foreign language tone words during initial stages of learning. *Journal of Cognitive Neuroscience* 34: 127–52. [\[CrossRef\]](#)
- Poe, Gina R., Stephen Foote, Oxana Eschenko, Joshua P. Johansen, Sebastien Bouret, Gary Aston-Jones, and Carolyn W. Harley. 2020. Locus coeruleus: A new look at the blue spot. *Nature Reviews Neuroscience* 21: 644–59. [\[CrossRef\]](#) [\[PubMed\]](#)
- Pruitt, David T., Ariel N. Schmid, Lily J. Kim, Caroline M. Abe, Jenny L. Trieu, Connie Choua, Seth A. Hays, Michael P. Kilgard, and Robert L. Rennaker. 2016. Vagus nerve stimulation delivered with motor training enhances recovery of function after traumatic brain injury. *Journal of Neurotrauma* 33: 871–79. [\[CrossRef\]](#)
- Quinkert, Amy Wells, Vivek Vimal, Zachary M. Weil, George N. Reeke, Nicholas D. Schiff, Jayanth R. Banavar, and Donald W. Pfaff. 2011. Quantitative descriptions of generalized arousal, an elementary function of the vertebrate brain. *Proceedings of the National Academy of Sciences of USA* 108: 15617–23. [\[CrossRef\]](#) [\[PubMed\]](#)
- Ranjbar-Slamloo, Yadollah, and Zeinab Fazlali. 2020. Dopamine and noradrenaline in the brain; overlapping or dissociate functions? *Frontiers in Molecular Neuroscience* 12: 334. [\[CrossRef\]](#)
- Raut, Ryan V., Abraham Z. Snyder, Anish Mitra, Dov Yellin, Naotaka Fujii, Rafael Malach, and Marcus E. Raichle. 2021. Global waves synchronize the brain's functional systems with fluctuating arousal. *Science Advances*. [\[CrossRef\]](#)
- Reetzke, Rachel, Zilong Xie, Fernando Llanos, and Bharath Chandrasekaran. 2018. Tracing the trajectory of sensory plasticity across different stages of speech learning in adulthood. *Current Biology* 28: 1419–27.e4. [\[CrossRef\]](#)
- Reimer, Jacob, Matthew J. McGinley, Yang Liu, Charles Rodenkirch, Qi Wang, David A. McCormick, and Andreas S. Tolias. 2016. Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. *Nature Communications* 7: 13289. [\[CrossRef\]](#) [\[PubMed\]](#)
- Remez, Robert E., Philip E. Rubin, David B. Pisoni, and Thomas D. Carrell. 1981. Speech perception without traditional speech cues. *Science* 212: 947–49. [\[CrossRef\]](#)
- Ruggiero, David A., Mark D. Underwood, Joseph John Mann, Muhammad Anwar, and Victoria Arango. 2000. The human nucleus of the solitary tract: Visceral pathways revealed with an “in vitro” postmortem tracing method. *Journal of the Autonomic Nervous System* 79: 181–90. [\[CrossRef\]](#)
- Sadaghiani, Sepideh, and Andreas Kleinschmidt. 2013. Functional interactions between intrinsic brain activity and behavior. *NeuroImage* 80: 379–86. [\[CrossRef\]](#)
- Sadakata, Makiko, and James M. McQueen. 2013. High stimulus variability in nonnative speech learning supports formation of abstract categories: Evidence from Japanese geminates. *The Journal of the Acoustical Society of America* 134: 1324–35. [\[CrossRef\]](#)
- Samuel, Arthur G. 1981. Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General* 110: 474–94. [\[CrossRef\]](#)
- Sara, Susan J. 2009. The locus coeruleus and noradrenergic modulation of cognition. *Nature Reviews Neuroscience* 10: 211–23. [\[CrossRef\]](#)
- Sara, Susan J., and Sebastien Bouret. 2012. Orienting and reorienting: The locus coeruleus mediates cognition through arousal. *Neuron* 76: 130–41. [\[CrossRef\]](#) [\[PubMed\]](#)

- Sara, Susan J., Andrey Vankov, and Anne Hervé. 1994. Locus coeruleus-evoked responses in behaving rats: A clue to the role of noradrenaline in memory. *Brain Research Bulletin* 35: 457–65. [\[CrossRef\]](#)
- Satpute, Ajay B., Philip A. Kragel, Lisa Feldman Barrett, Tor D. Wager, and Marta Bianciardi. 2019. Deconstructing arousal into wakeful, autonomic and affective varieties. *Neuroscience Letters* 693: 19–28. [\[CrossRef\]](#) [\[PubMed\]](#)
- Scharinger, Mathias, Molly J. Henry, and Jonas Obleser. 2013. Prior experience with negative spectral correlations promotes information integration during auditory category learning. *Memory & Cognition* 41: 752–68. [\[CrossRef\]](#)
- Schevernels, Hanne, Marlies E. van Bochove, Leen De Taeye, Klaas Bombeke, Kristl Vonck, Dirk Van Roost, Veerle De Herdt, Patrick Santens, Robrecht Raedt, and C. Nico Boehler. 2016. The effect of vagus nerve stimulation on response inhibition. *Epilepsy & Behavior* 64: 171–79. [\[CrossRef\]](#)
- Schuerman, William L., Kirill V. Nourski, Ariane E. Rhone, Matthew A. Howard, Edward F. Chang, and Matthew K. Leonard. 2021. Human intracranial recordings reveal distinct cortical activity patterns during invasive and non-invasive vagus nerve stimulation. *Scientific Reports* 11: 1–14. [\[CrossRef\]](#) [\[PubMed\]](#)
- Scott, Sophie K., and Carolyn McGettigan. 2013. The neural processing of masked speech. *Hearing Research* 303: 58–66. [\[CrossRef\]](#)
- Sharon, Omer, Firas Fahoum, and Yuval Nir. 2021. Transcutaneous vagus nerve stimulation in humans induces pupil dilation and attenuates alpha oscillations. *The Journal of Neuroscience* 41: 320–30. [\[CrossRef\]](#) [\[PubMed\]](#)
- Shetake, Jai A., Navzer D. Engineer, Will A. Vrana, Jordan T. Wolf, and Michael P. Kilgard. 2012. Pairing tone trains with vagus nerve stimulation induces temporal plasticity in auditory cortex. *Experimental Neurology* 233: 342–49. [\[CrossRef\]](#)
- Sohoglu, Ediz, Jonathan E. Peelle, Robert P. Carlyon, and Matthew H. Davis. 2012. Predictive top-down integration of prior knowledge during speech perception. *Journal of Neuroscience* 32: 8443–53. [\[CrossRef\]](#)
- Stanners, Robert F., Michelle Coulter, Allen W. Sweet, and Philip Murphy. 1979. The pupillary response as an indicator of arousal and cognition. *Motivation and Emotion* 3: 319–40. [\[CrossRef\]](#)
- Stein, Richard B., E. Roderich Gossen, and Kelvin E. Jones. 2005. Neuronal variability: Noise or part of the signal? *Nature Reviews Neuroscience* 6: 389–97. [\[CrossRef\]](#)
- Steriade, Mircea, Igor Timofeev, and François Grenier. 2001. Natural waking and sleep states: A view from inside neocortical neurons. *Journal of Neurophysiology* 85: 1969–85. [\[CrossRef\]](#)
- Strauss, Antje, Molly J. Henry, Mathias Scharinger, and Jonas Obleser. 2015. Alpha phase determines successful lexical decision in noise. *Journal of Neuroscience* 35: 3256–62. [\[CrossRef\]](#) [\[PubMed\]](#)
- Symmes, David, and Kenneth V. Anderson. 1967. Reticular modulation of higher auditory centers in monkey. *Experimental Neurology* 18: 161–76. [\[CrossRef\]](#)
- Taghia, Jalil, Weidong Cai, Srikanth Ryali, John Kochalka, Jonathan Nicholas, Tianwen Chen, and Vinod Menon. 2018. Uncovering hidden brain state dynamics that regulate performance and decision-making during cognition. *Nature Communications* 9: 2505. [\[CrossRef\]](#)
- Unsworth, Nash, and Matthew K. Robison. 2017. A locus coeruleus-norepinephrine account of individual differences in working memory capacity and attention control. *Psychonomic Bulletin & Review* 24: 1282–311. [\[CrossRef\]](#)
- Urbán, Michael A., Charles W. Lafe, Tyler W. Simpson, George F. Wittenberg, Bharath Chandrasekaran, and Douglas J. Weber. 2021. Electrical stimulation of the external ear acutely activates noradrenergic mechanisms in humans. *Brain Stimulation* 14: 990–1001. [\[CrossRef\]](#)
- Van Berkum, Jos J. A., Colin M. Brown, Pienie Zwitserlood, Valesca Kooijman, and Peter Hagoort. 2005. Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31: 443–67. [\[CrossRef\]](#)
- Van Bockstaele, Elisabeth J., James Peoples, and Patti Telegan. 1999. Efferent projections of the nucleus of the solitary tract to peri-locus coeruleus dendrites in rat brain: Evidence for a monosynaptic pathway. *Journal of Comparative Neurology* 412: 410–28. [\[CrossRef\]](#)
- Van Lysebettens, Wouter, Kristl Vonck, Lars Emil Larsen, Latoya Stevens, Charlotte Bouckaert, Charlotte Germonpré, and Mathieu Sprengers. 2020. Identification of vagus nerve stimulation parameters affecting rat hippocampal electrophysiology without temperature effects. *Brain Stimulation* 13: 1198–206. [\[CrossRef\]](#) [\[PubMed\]](#)
- Ventureyra, Enrique C. G. 2000. Transcutaneous vagus nerve stimulation for partial onset seizure therapy. *Child's Nervous System* 16: 101–12. [\[CrossRef\]](#) [\[PubMed\]](#)
- Vidaurre, Diego, Stephen M. Smith, and Mark W. Woolrich. 2017. Brain network dynamics are hierarchically organized in time. *Proceedings of the National Academy of Sciences of USA* 114: 12827–32. [\[CrossRef\]](#)
- Vonck, Kristl E. J., and Lars E. Larsen. 2018. Vagus Nerve Stimulation. In *Neuromodulation*. Amsterdam: Elsevier, pp. 211–20. [\[CrossRef\]](#)
- Vonck, Kristl, Robrecht Raedt, Joke Naulaerts, Frederick De Vogelaere, Evert Thiery, Dirk Van Roost, Bert Aldenkamp, Marijke Miatton, and Paul Boon. 2014. Vagus nerve stimulation . . . 25 years later! What do we know about the effects on cognition? *Neuroscience & Biobehavioral Reviews* 45: 63–71. [\[CrossRef\]](#)
- Warren, Richard M. 1970. Perceptual restoration of missing speech sounds. *Science* 167: 392–93. [\[CrossRef\]](#)
- Warren, Christopher M., Klodiana D. Tona, Lineke Ouwerkerk, Jeroen van Paridon, Fenna Poletiek, Henk van Steenbergen, Jos A. Bosch, and Sander Nieuwenhuis. 2019. The neuromodulatory and hormonal effects of transcutaneous vagus nerve stimulation as evidenced by salivary alpha amylase, salivary cortisol, pupil diameter, and the P3 event-related potential. *Brain Stimulation* 12: 635–42. [\[CrossRef\]](#) [\[PubMed\]](#)

- Waschke, Leonhard, Sarah Tune, and Jonas Obleser. 2019. Local cortical desynchronization and pupil-linked arousal differentially shape brain states for optimal sensory performance. *eLife* 8: e51501. [\[CrossRef\]](#)
- Werker, Janet F., and Takao K. Hensch. 2015. Critical periods in speech perception: New directions. *Annual Review of Psychology* 66: 173–96. [\[CrossRef\]](#) [\[PubMed\]](#)
- Whyte, John. 1992. Attention and arousal: Basic science aspects. *Archives of Physical Medicine and Rehabilitation* 73: 940–49. [\[CrossRef\]](#) [\[PubMed\]](#)
- Yakunina, Natalia, Sam Soo Kim, and Eui-Cheol Nam. 2017. Optimization of transcutaneous vagus nerve stimulation using functional MRI. *Neuromodulation: Technology at the Neural Interface* 20: 290–300. [\[CrossRef\]](#) [\[PubMed\]](#)
- Yap, Jonathan Y. Y., Charlotte Keatch, Elisabeth Lambert, Will Woods, Paul R. Stoddart, and Tatiana Kameneva. 2020. Critical review of transcutaneous vagus nerve stimulation: Challenges for translation to clinical practice. *Frontiers in Neuroscience* 14: 284. [\[CrossRef\]](#)
- Yi, Han Gyo, and Bharath Chandrasekaran. 2016. Auditory categories with separable decision boundaries are learned faster with full feedback than with minimal feedback. *The Journal of the Acoustical Society of America* 140: 1332–35. [\[CrossRef\]](#)
- Yi, Han Gyo, Bharath Chandrasekaran, Kirill V. Nourski, Ariane E. Rhone, William L. Schuerman, Matthew A. Howard, Edward F. Chang, and Matthew K. Leonard. 2021. Learning nonnative speech sounds changes local encoding in the adult human cortex. *Proceedings of the National Academy of Sciences of USA* 118: e2101777118. [\[CrossRef\]](#)
- Yi, Han Gyo, Matthew K. Leonard, and Edward F. Chang. 2019. The encoding of speech sounds in the superior temporal gyrus. *Neuron* 102: 1096–110. [\[CrossRef\]](#)
- Yi, Han Gyo, W. Todd Maddox, Jeanette A. Mumford, and Bharath Chandrasekaran. 2016. The role of corticostriatal systems in speech category learning. *Cerebral Cortex* 26: 1409–20. [\[CrossRef\]](#) [\[PubMed\]](#)
- Young, Christina B., Gal Raz, Daphne Everaerd, Christian F. Beckmann, Indira Tendolkar, Talma Hendler, Guillén Fernández, and Erno J. Hermans. 2017. Dynamic shifts in large-scale brain network balance as a function of arousal. *Journal of Neuroscience* 37: 281–90. [\[CrossRef\]](#)
- Yu, Luodi, and Yang Zhang. 2018. Testing native language neural commitment at the brainstem level: A cross-linguistic investigation of the association between frequency-following response and speech perception. *Neuropsychologia* 109: 140–48. [\[CrossRef\]](#) [\[PubMed\]](#)
- Zekveld, Adriana A., Thomas Koelewijn, and Sophia E. Kramer. 2018. The pupil dilation response to auditory stimuli: Current state of knowledge. *Trends in Hearing* 22: 1–25. [\[CrossRef\]](#)
- Zhang, Yang, Patricia K. Kuhl, Toshiaki Imada, Makoto Kotani, and Yoh'ichi Tohkura. 2005. Effects of language experience: Neural commitment to language-specific auditory patterns. *NeuroImage* 26: 703–20. [\[CrossRef\]](#) [\[PubMed\]](#)
- Zhang, Yang, Patricia K. Kuhl, Toshiaki Imada, Paul Iverson, John Pruitt, Erica B. Stevens, Masaki Kawakatsu, Yoh'ichi Tohkura, and Iku Nemoto. 2009. Neural signatures of phonetic learning in adulthood: A magnetoencephalography study. *NeuroImage* 46: 226–40. [\[CrossRef\]](#)