Temporal Course of Visual Modulation in Cat Auditory Systems

Xiaohan Bao

Integrated Program in Neuroscience

McGill University, Montreal, Canada

May 2024

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Ph.D. in Neuroscience

Table of Content

Abstract

Sensory system in human, as well as in many other species, encodes the external world through multiple modalities. Signals emitted when an event occurs are integrated by the brain to form a multisensory perception. When the pathway of one sensory modality is damaged, the brain will adapt to coping with new situation using the rest sensory modalities. Such cross-modal interaction must involve interconnections between different sensory systems. Examples of audiovisual interaction are commonly experienced such as when one hears a speech better while seeing the lip moving. People with blindness or deafness often demonstrated supranormal hearing or visual functions. Understanding the mechanism of neural plasticity underlying audiovisual interaction is crucial for advancing our current medical treatment to diseases of sensory loss, such as congenital deafness. My thesis investigated the role of the auditory cortex in the function of audiovisual processing, based on well-established cat model. Previously, it has been demonstrated that cat auditory cortex was extensively involved in visual functions. Using a novel paradigm of audiovisual stimulus and recording paradigms, I confirmed the visual modulation of auditory responses measured by auditory evoked potentials (AEPs) derived from scalp-recorded EEG in cats lightly anesthetized by dexmedetomidine (see Chapter 3) and click elicited local field potentials (LFPs) in awake primary auditory cortex (A1) (see Chapter 5). The association between audiovisual temporal disparity and the amount of visual modulation was also characterized. The results revealed visual suppression of click LFPs and AEPs for a visual-to-auditory delay of about 100 milliseconds, which was consistent with previous findings. Changes in click LFPs and AEPs, however, were also discovered with longer visual-to-auditory delays, suggesting a recursive or maybe periodic pattern in the temporal course of visual modulation. Taken together, these

experiments have made successful attempts on new research methodology and expanded the existing knowledge of audiovisual processing in cat auditory cortex.

Résumé

Le système sensoriel humain, ainsi que chez de nombreuses autres espèces, encode le monde extérieur à travers de multiples modalités. Les signaux émis lorsqu'un événement survient sont intégrés par le cerveau pour former une perception multisensorielle. Lorsque la voie d'une modalité sensorielle est endommagée, le cerveau s'adaptera pour faire face à une nouvelle situation en utilisant les autres modalités sensorielles. Une telle interaction inter-modale doit impliquer des interconnexions entre différents systèmes sensoriels. Des exemples d'interactions audiovisuelles sont couramment vécues, par exemple lorsque l'on entend mieux un discours en voyant les lèvres de celui qui parle bouger. Souvent, les personnes atteintes de cécité ou de surdité présentent respectivement une audition ou des fonctions visuelles supérieures à la moyenne. Comprendre les mécanismes de plasticité neuronale sous-jacents à l'interaction audiovisuelle est crucial pour faire progresser le traitement médical actuel des maladies de perte sensorielle telle que la surdité congénitale. Ma thèse porte sur le rôle du cortex auditif dans le fonctionnement du traitement audiovisuel, en utilisant le chat comme espèce modèle bien établie. Auparavant, il a été démontré que le cortex auditif du chat était largement impliqué dans les fonctions visuelles. En utilisant un nouveau paradigme de stimulation et d'enregistrement audiovisuels, j'ai confirmé la modulation visuelle des potentiels évoqués auditifs (AEP) dérivés de l'EEG enregistré chez des chats légèrement anesthésiés sous dexmédétomidine (voir chapitre 3) et des potentiels de champ locaux induits par des clics (click-LFP) dans le cortex auditif primaire (A1) chez des chats éveillés (voir chapitre 5). L'association entre la disparité temporelle audiovisuelle et la quantité de modulation visuelle a également été caractérisée. Les résultats ont révélé une suppression visuelle des click-LFP et des AEP pour un retard visuel-vers-auditif d'environ 100 millisecondes, ce qui était cohérent avec les découvertes précédentes. Des changements dans les click-LFP et les AEP ont

également été observés avec des retards plus longs, suggérant un motif récursif ou peut être périodique dans le décours temporel de la modulation visuelle. Ensemble, ces expériences ont permis d'élaborer une nouvelle méthodologie de recherche et élargi les connaissances existantes sur le traitement audiovisuel dans le cortex auditif du chat.

Acknowledgement

First, I would like to express my sincere gratitude to my supervisor, Dr. Stephen Lomber. It is my greatest fortunate to have his guidance and support as my Ph.D. supervisor throughout the years. All the challenge attached to my experimental works was made effortless, thanks to his extensive hands-on experience and exquisite skills for neurosurgical procedures. Without his expertise, none of the research in this thesis would have taken place. Most importantly, I learnt from him not just about doing scientific work, but also how to make wise decisions when facing challenges and difficulties, and how to navigate through obstacles in front of my goals. Among all the scholars I have met, there is not a second person like him, who has taught me so much on what it is like for taking research as a life-long career, with every question fueled by inquisitive curiosity and every experiment carried out by unshakeable determination.

I would also like to thank other lab members, including Dr. Carmon Wong for her advice in chronic recording techniques, Stephen Gordon for developing the apparatus of dispensing cat treats, and Anas Salloum for collecting some of the behavioral data. I can't express my appreciation enough to all my lab mates, Alessandra Sacco, Carina Sabourin, Charlotte Kruger, Simon Zhu , Danial Ghiaseddin, Mahdi Mahdavi, Hadi Parvin, Gabe Wooten-Soto, Sena Tavukçu, and Thomas Mitzelfelt. There have been uncountable moments when I feel deeply touched by their supports and accompany through this journey.

Also, my research projects could never be carried on without the valuable help from the veterinary staff, including Paisley Barnes, Karen Hope, Dr. Jessica D'Amico, and Dr. Julie Chevrette. In particular, I would like to thank our lab technician, Paisley Barnes, for assisting all the implant surgeries and administrating animal anesthesia for all the procedures.

I would also like to thank my advisory committee at McGill University, who are Dr. Curtis Baker, Dr. Sarah Wolley, and Dr. Erik Cook, as well as my advisory committee at Western University, who are Dr. Brian Allman, Dr. Brian Corneil, and Dr. Ewan Macpherson. I will not be able to get towards the completion of my thesis without all the advice and suggestions you have given to me generously through each of the committee meeting.

Finally, I would like to thank my parents and friends, who have kept me accompanied and supported throughout the years. The past couple of years during the pandemic have been especially challenging for all of us. To my friends, Jiayi, Yuli, Jing, Ziqi, Wenji, and Shuo, thank you for being like my family during the years when we can't travel back to our country. Special thanks to Dr. Clementine Bodin for editing the French abstract of this thesis. And to my parents, thank you for the phone calls and messages that have been making me feel you are by my sides all the time.

Contribution to Original Knowledge

The present thesis established a restrain-free extracellular recording protocol in awake cats and designed a novel audiovisual stimulus paradigm to probe the effect of audiovisual temporal disparity on visual modulation of auditory excitability. Compared to other models (e.g., human, non-human primates or rodents), cats were rarely used for studying the functional mechanism of multisensory interaction. Therefore, this thesis greatly contributes to the literature on this subject. In the first study, I described the behavioral training process and demonstrated that five cats successfully acquired the fixation task to an extent sufficient for being tested in a psychophysical experiment. The protocol and experiment devices are expected to be used as a standard using for behavioral and electrophysiological studies in awake cats. In the second study, I characterized the time course of visual modulation as a function of audiovisual temporal disparity, through extracellular recording from chronically-implanted electrode matrices in the primary auditory cortex of two cat subjects. The audiovisual stimulus paradigm designed for this study allowed sampling of stimulus delays that are a couple times more than previous studies in a more efficient way. It was demonstrated that complex multi-peak delay functions of audio-tactile interaction shown in macaque also present in cats for audiovisual interaction. The pattern of delay dependency reveals the features of neural circuits in sensory cortices, implying the integrative functions of the primary auditory cortex for ascending auditory inputs and lateral / descending visual inputs from other brain regions.

To explore the effect of audiovisual disparity at the level of neural population, I turned to EEG recording in the second half of my thesis. Although event-related potentials (ERPs) technique has been widely used as a neurophysiological approach in human subjects, it is much less popular

in animal model. In the third study, I exhibited the differential suppressive effect of isoflurane between auditory and visually evoked potentials. These findings informed future animal EEG studies that cautions must be taken selecting isoflurane as anesthetics, as it may take effect unequally among different sensory modalities. In the last study, I characterized the effect of audiovisual disparity on auditory evoked potentials in thirteen cats, with the stimulus paradigm used in the second study. The outcome of this study suggests that the pattern of delay dependency can be characterized with scalp-recorded EEG signal, which reflects the aggregated electrical activities time-locked to stimulus inputs in related sensory areas such like the primary auditory cortex.

Altogether, these results demonstrated that temporal course of visual modulation in auditory cortex can be characterized with various electrophysiological techniques in cats as an animal model. Previous studies have also shown behavioral and electrophysiological evidence for cortical re-organization following deafness in cats. With continued investigation, our understanding in cat cortex on the subject of cross-modal interaction will be further advanced and inspire the upgrade of neuroprosthetic devices (e.g., cochlear implant) in treating sensory deprivation.

List of Figures and Tables

List of Abbreviation

AAF	anterior auditory cortex
A1	primary auditory cortex
ABR	auditory brain-stem response
AEP	auditory evoked potential
AES	anterior ectosylvian sulcus
CI	cochlear implant
CN	cochlear nucleus
DZ	dorsal field of auditory cortex
ERPs	event-related potentials
fMRI	functional magnetic resonance imaging
fAES	the auditory field of anterior ectosylvian sulcus
IC	inferior colliculus
LFP	local field potential
MGN	medial genicular nucleus
OFC	orbital frontal cortex
PAF	posterior auditory field
PET	positron emission tomography
RT	Reaction time
SC	Superior colliculus
SOC	superior olivary complex
STG	superior temporal gyrus
STS	superior temporal sulcus
VEP	visually evoked potential

1. Chapter 1: General Introduction

1.1. Overview

Neural mechanism underlying cross-modal interaction has been extensively studied (Bavelier & Neville, 2002; A. K. Lee, Wallace, Coffin, Popper, & Fay, 2019). After explaining the relationship between cross-modal processing and cross-modal plasticity (**1.2.1**), I will give a brief review on some classical neurophysiological studies that demonstrated the instantaneous neuronal integration of stimuli from multiple modalities and its general principles (**1.2.2**), and how chronically depriving one sensory modality can affect other sensory modalities (**1.2.3**). Sensory deprivation is no longer irreversible due to the advancement in medicine, and therefore our research is also motivated ultimately by the need of understanding on how the multisensory system of the brain can be harnessed in different clinical settings (**1.2.4**), such as auditory restoration through cochlear implant.

Cross-modal interactions that occur along a traditional uni-sensory pathway instead of a polymodal area have received more and more research attention in the past two decades (Kayser, 2010; Kayser & Logothetis, 2007; Murray et al., 2016). Particularly, the idea that primary sensory cortices are strictly uni-sensory has been questioned. Here, I will introduce some discoveries of cross-modal effects in the auditory system, including the subcortical auditory areas (**1.3.1**) and the auditory cortex (**1.3.2**) both. Although the latest research seems to favor cerebral cortex for its accessibility in both human and animal models, it is still technically challenging for differentiating the cross-modal effects that truly emerge from cortical inter-connections from those carried-over subcortically.

The interplay of auditory and visual processing has a unique position among multisensory research, probably due to the use of auditory and visual speech by human. Moreoever, both auditory and visual sensations can carry information emitted from a larger distance from the receiving individuals than tactile sense. In particular, a number of studies have shown that visual information can influence auditory processing, and it involves neural modulation of the auditory cortex by visual stimulus (**1.4.1**). On the other hand, hearing loss is known to trigger both anatomic and/or functional conversions of the auditory system to other sensory modalities (**1.4.2**). It has been proposed that the audiovisual interaction can be manipulated by attention and may be indicative for normal and abnormal cognitive development (**1.4.3**). To generate new knowledge that may make theoretical contribution in the questions above, my thesis is staged to study visual modulation of auditory processing with a specific target on the auditory cortex.

Although a vast number of studies have demonstrated the behavioral outcomes of the temporal disparity in audiovisual stimuli, our understanding of its neurophysiological correlates is still very limited especially in the auditory cortex, despite that the temporal integration of audiovisual stimuli has been put forward for more than 30 years (**1.5.1**). The effect of audiovisual temporal disparity is often studied in the perspective of audiovisual temporal processing, where subjects were asked or trained to make an explicit decision on the temporal order or simultaneity of a pair of auditory and visual stimuli. (**1.5.2**). However, the temporal disparity can also affect how visual (auditory) stimulus modulates auditory (visual) processing implicitly, where the audiovisual asynchrony is task-irrelevant (**1.5.3**).

On the background of previous research outlined above, the present thesis was focused on establishing extracellular recording from awake auditory cortex and scalp-recorded EEG in cats to bridge the most recent findings on audiovisual interactions from human and other animal models.

Audiovisual disparity was selected as the major factor to investigate, because of its popularity among psychophysics studies but elusive neural mechanism (**1.6**).

1.2. Cross-modal interaction as a general feature of sensory system

1.2.1. Cross-modal plasticity and multisensory processing

Interactions between different sensory modalities have been observed in human (Gao et al., 2023; Lessard, Paré, Lepore, & Lassonde, 1998; Macaluso & Driver, 2005) and animals (Cloke, Jacklin, & Winters, 2015; M Alex Meredith, Nemitz, & Stein, 1987; Musacchia & Schroeder, 2009). Stimuli for different sensory modalities are transmitted through different types of media and can be independent in the physical world. Once the stimuli are received by the central nerve system and encoded as in neural signals, they are still processed in parallel in general but with complex inter-connections. Consequently, stimulus from one modality can instantaneously interfere the neural processing in another modality (Spence, Senkowski, & Röder, 2009), and furthermore, the absence of stimulus inputs to one modality over a long time can permanently alter another modality functionally and anatomically (Bavelier & Neville, 2002). While "multisensory processing" and "cross-modal plasticity" have been used to describe the instantaneous and the permanent cross-modal interactions, it is likely that, in terms of the neurophysiological foundations, they are just the two sides of the same coin (Alais, Newell, & Mamassian, 2010; M Alex Meredith & Lomber, 2017).

Having multiple sensory modalities has apparent advantages over solely relying on one sensory modality. It has been tested in human and many animal species that input from additional sensory modalities improves behavioral outcomes, such as higher performance in detecting the object (McDonald, Teder-Salejarvi, & Hillyard, 2000), and higher speech intelligibility (Sumby & Pollack, 1954). A classic framework of multisensory benefit, known as "race model" inequality

(Miller, 1982; Miller, 1986), showed that redundant information increases the speed of task performance. The model proposed that the probability distribution of reaction times (RTs) in response to multisensory stimuli is in violation against the summation of RT probability distributions estimated using unisensory stimuli. This has been evidenced not only in human but also animal models, including non-human primates (Cappe, Murray, Barone, & Rouiller, 2010; Fetsch, Turner, DeAngelis, & Angelaki, 2009; Lessard et al., 1998), ferrets (Hammond-Kenny, Bajo, King, & Nodal, 2017), cats (Gingras, Rowland, & Stein, 2009), and rodents (Hirokawa et al., 2011; Meijer, Pie, Dolman, Pennartz, & Lansink, 2018).

On the other hand, bias and error can also derive from multisensory stimuli under certain circumstances. Multisensory stimuli with conflict information can cause mistakes in both lower-level perceptual outcomes, e.g., identifying spatial information as in Ventriloquist effect (Alais, 2004), and higher-level cognitive outcomes, e.g., recognizing/discriminating consonants as in McGurk effect (McGurk, 1976). Such mistakes can be specific to audiovisual conflict as shown in both examples, or to conflicts between other sensory modalities. For example, rubber hand illusion is due to conflict of visual, tactile, and proprioception information (Botvinick, 1998). Recently, it was found that vestibular stimulus can bias the perception of visual stimulus as well (Bai, 2020).

1.2.2. Multisensory brain areas (superior colliculus and superior temporal cortex)

Multisensory functions are first localized in brain regions exclusively different from those traditionally known as unisensory areas (Schepers, Hipp, Schneider, Röder, & Engel, 2012). In the current perspective, this is no longer the case. Recent findings showed that, unisensory areas are, to some extent, involved in cross-modal interaction (Bizley, Nodal, Bajo, Nelken, & King, 2007;

Kayser, Petkov, Augath, & Logothetis, 2005, 2007; Kayser, Petkov, & Logothetis, 2008; Lakatos, Chen, O'Connell, Mills, & Schroeder, 2007; M Alex Meredith & Allman, 2015; Noesselt et al., 2010). That being said, our knowledge about how brain processes multisensory inputs is largely influenced by the findings in those multisensory brain regions discovered in those early days. Before the introduction to the multisensory properties in the auditory system, it is essential to briefly review the literatures on these "multisensory" brain regions.

The earliest most influential neurophysiology studies on multisensory processing were established in cat superior colliculus (SC) with extracellular recordings (Meredith et al., 1987; Meredith et al., 1986, 1996; Stein et al., 1993; Wallace et al., 1997). Multisensory interactions in SC neurons were identified by comparing firing rates in the responses elicited by two or three unisensory stimuli (e.g., auditory, visual, or tactile stimulus present alone) with the response elicited by multisensory stimuli (e.g., auditory and visual stimuli co-present). It was proposed by this line of work that there are a few "rules" that determine the behavior of a multisensory SC neuron: (1) Spatial factor: response to multisensory stimuli combined can be enhanced or suppressed only when the stimuli are co-localized in space. (2) Temporal factor: response to multisensory stimuli combined can be enhanced or suppressed when the inter-stimulus delay falls within a certain temporal window. (3) Multiplicative magnitude: the enhancement of a multisensory combination can go beyond the magnitude of a simple summation of unisensory responses. (4) Inverse effectiveness: the enhancement of multisensory combination is larger when its unisensory stimuli are less effective by themselves.

Meanwhile, cortical processing of multisensory stimuli has been unveiled in different species. One study, also in cats, showed that SC neurons remain multi-modal responsive but does not demonstrate nonlinear enhancement of multisensory interaction without the functional inputs

from the anterior ectosylvian sulcus (AES) (Alvarado, Stanford, Rowland, Vaughan, & Stein, 2009). In macaque, Benevento (1977) recorded single neurons from superior temporal sulcus (STS) and orbital frontal cortex (OFC) that can be excited by both visual and auditory stimuli. The involvement of STS in human audiovisual processing of speech was also confirmed with positron emission tomography (PET) (Macaluso, George, Dolan, Spence, & Driver, 2004) and functional magnetic resonance imaging (fMRI) (Macaluso et al., 2004; Stevenson, Altieri, Kim, Pisoni, & James, 2010; Stevenson, Kim, & James, 2009; Stevenson, VanDerKlok, Pisoni, & James, 2011), as well as the manipulation of transcranial magnetic stimulation (TMS) (Beauchamp, Nath, & Pasalar, 2010; Nath & Beauchamp, 2012). In addition, O'Dhaniel (2005) found neurons in macaque lateral and medial intraparietal areas (LIP and MIP) responding to auditory and visual target for saccadic movements.

As mentioned at the beginning of this section, recent reviews have challenged the traditional view about the exclusive unisensory and multisensory divisions of brain areas, as studies have demonstrated anatomic and functional evidence for cross-modal interactions at primary sensory cortices. In short, the understanding of neural mechanism underlying multisensory processing is still advancing. Particularly, experiments studying the emerging cross-modal cortical reorganization after sensory deprivation as a special form of multisensory interaction (i.e., a misplaced functional allocation), which will be introduced in the next section, may provide its unique insights.

1.2.3. Sensory loss and brain plasticity

When one sensory modality is deprived for an individual, it becomes particularly beneficial if the remaining sensory modalities can be further exploited. Taking deafness as an example, it has been reviewed (Alencar, Butler, & Lomber, 2019) that deaf individuals were better at peripheral vision (Codina, Buckley, Port, & Pascalis, 2011; Stephen G Lomber, M Alex Meredith, & Andrej Kral, 2010), detection of visual motion (Scott, Karns, Dow, Stevens, & Neville, 2014; Stevens & Neville, 2006), detection of the change in frequency of tactile vibration (Levänen & Hamdorf, 2001), discrimination in object and facial changes (Megreya and Bindemann 2017). They also showed better visuomotor synchronization (Iversen et al., 2015). Likewise, it has been reported that blind individuals were better at detecting a deviant sound among standard (Röder, Rösler, & Neville, 1999), had better spectral and temporal resolution (Shim, Go, Lee, Choi, & Won, 2019), and sound categorizations (Schepers et al., 2012).

Cats are well-known as an animal model in studying cross-modal plasticity following sensory deprivation. A study in early 90s showed that visually-deprived cats demonstrated better performance in sound localization (J. P. Rauschecker & Kniepert, 1994) especially for sounds coming from the positions lateral or rear to the subjects. Consistent with this behavioral finding, extracellular recordings in superior colliculi of visually deprived cats discovered a large increase of auditory-responsive neurons compared to sighted controls (J. Rauschecker & Harris, 1983). Similarly, congenitally deaf cats demonstrated improved performance in visual localization and movement detection tasks (Stephen G Lomber et al., 2010), and these behavioral improvements required the participation of posterior auditory fields (PAF) and dorsal field of auditory cortex (DZ), respectively.

Additionally, in people with blindness (Bhattacharjee, Amanda, Lisak, Vargas, & Goldreich, 2010) and deafness (Levänen & Hamdorf, 2001), there are studies showing improved somatosensory processing. In support of such observations, animal studies have shown increased somatosensory activities in deaf auditory cortex as well (Allman, Keniston, & Meredith, 2009; M Alex Meredith & Lomber, 2011). The evidence reviewed here altogether suggests a highly plastic sensory system capable of compensating a lost sensory modality with remaining sensory modalities. More importantly, the anatomic architecture of cross-modal plasticity may be shared with those underlying multisensory processing.

1.2.4. Maladaptive brain plasticity impedes outcome of hearing restoration

Hearing impairment is the second most prevalent impairment (following anaemia) as an outcome of diseases, affecting 1.33 billion (18.5%) people in the world (Vos et al., 2016). Untreated hearing impairment significantly undermines cognition (Taljaard, Olaithe, Brennan-Jones, Eikelboom, & Bucks, 2016) especially for prelingual hearing impairment caused genetically. Cochlear prosthetics (cochlear implants, CIs) are an effective technique to replace the cochlea not functioning. The replacement of damaged cochlea by itself, however, cannot revert brain reorganization occurring during the period of auditory deprivation.

More than 20 years ago, a case study, using positron emission tomography (PET) technique, reported visual activations in the secondary auditory cortex by sign language and auditory activations in the primary auditory cortex through cochlear implant (Nishimura et al., 1999). Since the sign-language activation was only observed in deaf but not hearing sign-language users, it was proposed by the authors that the absence of auditory activation in the secondary

auditory cortex may cause the ultimate failure of speech comprehension in this subject (Nishimura et al., 1999). In another word, the visual function of the auditory cortex obtained by cross-modal plasticity adapting to deafness may impede adapting to the electric hearing using cochlear implant, which makes such cross-modal plasticity a maladaptive one.

In support of this hypothesis, another PET study found the glucose metabolism in the auditory cortex before cochlear implantation was negatively correlated with the score of speech skill with cochlear implants (D. S. Lee et al., 2001). Similarly, visual activation in the auditory cortex, measured with visually evoked potentials post-implant, was also negatively correlated with speech intelligibility (Sandmann et al., 2012). At behavioral level, increase in visual performance (i.e., decreased reaction time) in a phonological task using orthographical information predicted worse speech comprehension in cochlear implant users (Lazard & Giraud, 2017).

In summary, although neural plasticity after hearing loss occurs in nature as an adaptive strategy of sensory system during abnormaldevelopment, such plasticity may be the barrier for cochlear implant to achieve its optimal outcome. Although neural plasticity after cochlear implant seems to adapt sensory system to electric hearing as well (Glennon, Svirsky, & Froemke, 2020), it is still unknown how much the normal trajectory of brain development for cognitive functions beyond auditory functions is compromised.

Therefore, to achieve optimal auditory restoration, it is important to understand the cross-modal plasticity in the auditory cortex by studying its functionality in both hearing and deaf subjects. In the next section, I will focus on existing literature about cross-modal interactions in the cortical and subcortical auditory pathway.

1.3. The central auditory system and how it is affected by non-auditory sensory inputs

1.3.1. The participation of sub-cortical auditory system in cross-modal interaction

Auditory system starts with cochlear hair cells modulating the firing of cochlear ganglion neurons (auditory nerves) (Middlebrooks, 2015), which make synapses with neurons at cochlear nucleus (CN). It is known that auditory responses in CN were influenced by somatosensory inputs in guinea pigs (Shore & Zhou, 2006), although whether CN receives visual inputs has not been tested yet. A majority of CN neurons project to the contralateral superior olivary complex (SOC) (Jennings, 2021). There is no literature showing SOC neurons receive inputs from non-auditory system, but a comparative study has shown that the size of medial superior olive was correlated with eye diameter across different species (Irving & Harrison, 1967). Similarly, it has been shown that lateral nucleus of the superior olive in blind mole rat, a subterranean rodent, was poorly differentiated (Bronchti, Heil, Scheich, & Wollberg, 1989). From ipsilateral SOC, as well as contralateral CN, auditory inputs were relayed to auditory midbrain inferior colliculus (IC). Evoked potential studies showed that auditory responses in IC neurons also were under modulation from somatosensory inputs in rats (Szczepaniak & Møller, 1993), from visual inputs in owls and mice (Bergan & Knudsen, 2009; Cheng, Fei, & Qu, 2019; Cheng, Guo, & Qu, 2020) and from eye positions in monkeys (Groh, Trause, Underhill, Clark, & Inati, 2001; Gruters & Groh, 2012). Ascending auditory neurons from IC project to medial genicular nucleus (MGN). A study showed that human MGN volume was increased after vision loss (Moro, Kelly, McKetton, Gallie, & Steeves, 2015). Responses in rat auditory thalamic neurons to acoustic inputs were also modulated by visual stimulus (Komura, Tamura, Uwano, Nishijo, & Ono, 2005). In addition to the studies focusing on subcortical auditory nuclei (see **Fig. 1** for summary), one study also found multisensory processing in basal ganglion neurons in anesthetized cats (Nagy, Eördegh, Paróczy,

Márkus, & Benedek, 2006). It was later found that tactile and visual inputs were integrated in mouse striatum as well (Reig & Silberberg, 2014).

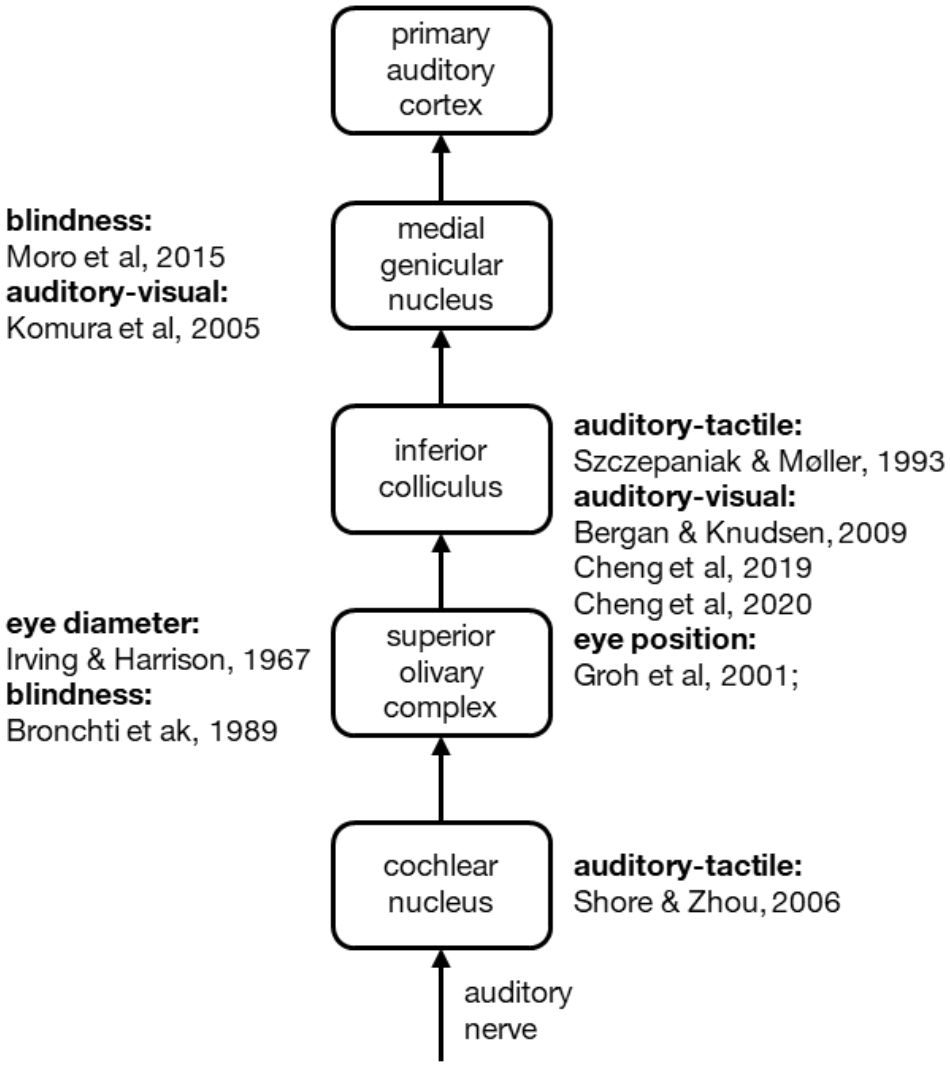


Figure 1. An illustration of ascending auditory pathway and its cross-modal interaction previously reported.

In human, two studies have investigated visual modulation of subcortical auditory pathway using electrophysiological approach (Caron-Desrochers, Schönwiesner, Focke, & Lehmann, 2018; Musacchia, Sams, Nicol, & Kraus, 2006). A suppression of the P0 wave from auditory middle-

latency responses by synchronized visual stimulus compared to non-synchronized visual stimulus was documented (Caron-Desrochers et al., 2018). This suppression, however, was not statistically significant. No visual modulation in amplitudes or latencies in wave 1 to 5 from auditory brain-stem responses (ABRs) was observed. This was however in contradiction with another study, where viewing lip movements significantly postponed the onset of frequency following responses to speech for ~1 ms (Musacchia et al., 2006).

It is believed that visual deprivation can rewire subcortical networks (Bavelier & Neville, 2002). In a species of congenitally blind mole rat, Spalax ehrenbergi, neurons from inferior colliculus were found to project to not only auditory but also visual thalamus and convert the occipital cortex to respond to auditory stimulus (Doron & Wollberg, 1994).

1.3.2. The participation of auditory cortex in cross-modal interaction

The evidence showing the involvement of the auditory cortex in cross-modal interaction has mainly come from studies using three techniques: (1) retrograde tracer, (2) neurophysiological recording, and (3) cooling loop deactivation. By combining all these techniques, it has been demonstrated that the auditory cortex is equipped with the anatomic foundations and functional relevance in visual and somatosensory processing. Studies using electrophysiology recordings have been carried out in various sub-regions of the auditory cortex in anesthetized ferrets (Bizley et al., 2007) and cats (Merrikhi, Kok, Carrasco, Meredith, & Lomber, 2022; Merrikhi, Kok, Lomber, & Meredith, 2023). This line of research has shown that neurons in auditory cortex are also influenced by somatosensory and visual stimulus. Particularly, in hearing impaired animals, an increase in cross-modal responses was observed.

Five sub-areas of the auditory cortex were investigated by the deposit of retrograde tracer and whole-brain histological analysis showed that they not only received projections from other auditory areas, but also visual areas and somatosensory areas (see **Table 1** and **Fig. 2**). In all five areas investigated (Butler, Chabot, & Lomber, 2016; Butler, de la Rua, Ward-Able, & Lomber, 2018; Chabot, Butler, & Lomber, 2015; Kok, Chabot, & Lomber, 2014; Wong, Chabot, Kok, & Lomber, 2015), the highest proportion of visual cortical inputs was found in the primary auditory cortex (A1) (10.8%) while the highest proportion of somatosensory cortical inputs was found in the anterior auditory field (AAF) (10%). Furthermore, such cross-modal projections may change in hearing and deaf individuals. For example, it was shown that projections from visual and somatosensory areas to AAF was significantly higher in deaf cats than normal hearing controls (Wong et al., 2015).

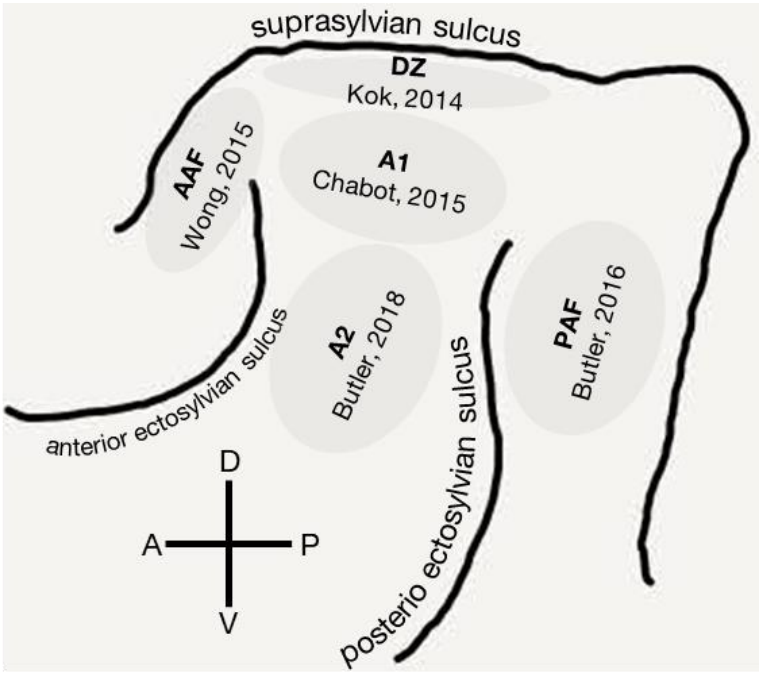


Figure 2. Subareas of cat auditory cortex where visual inputs have been previously investigated.

Table 1: Retrograde tracer studies revealing cross-modal projections in cat auditory cortex

Injected Area	Labeled cell in cortices tagged to sensory modalities			Citations
	Auditory	Visual	Somatosensory	
PAF	66%	2.2%	2.1%	Butler 2016
A1	46.5%	10.8%	0.7%	Chabot 2015
A2	56.3%	9.5%	2.4%	Butler 2018
DZ(only cortical)	46.3%	45.6%	0.6%	Kok 2014
AAF	53.9%	2.1%	10.0%	Wong 2015

Studies using neurophysiological/electrophysiological techniques record and measure the activities in the auditory cortex related to non-auditory stimuli. A great amount of human studies have revealed cross-modal cortical activation or modulation in the auditory cortex by visual (Finney, Clementz, Hickok, & Dobkins, 2003; Finney, Fine, & Dobkins, 2001) and tactile stimuli (Hoefer et al., 2013; Martuzzi et al., 2007; Nordmark, Pruszynski, & Johansson, 2012; Schürmann,

Caetano, Hlushchuk, Jousmäki, & Hari, 2006). In non-human primates, neurons in the primary auditory cortex responding to visual (Ghazanfar, Maier, Hoffman, & Logothetis, 2005; Kayser et al., 2008; Kayser, Petkov, & Logothetis, 2009) and tactile stimuli (Lakatos et al., 2007) have been discovered as well. In smaller animals, visual and tactile responses were also found in the auditory cortex of hearing and deaf animals (mouse: Hunt, Yamoah, & Krubitzer, 2006; ferret: M Alex Meredith & Allman, 2015; cat: M Alex Meredith & Lomber, 2011). Overall, deaf subjects demonstrated more visual or tactile activations in the auditory cortex.

Finally, the behavioral relevance of a sub-area of the auditory cortex can be studied with cooling loop deactivation. In deaf cats, deactivating sub-regions of the auditory cortex (posterior auditory field, PAF or dorsal zone, DZ) brought visual task performance (peripheral localization or movement detection, respectively) of deaf cats down to the levels observed in normal hearing cats (S. G. Lomber, M. A. Meredith, & A. Kral, 2010). Deactivating auditory field of the anterior ectosylvian sulcus (fAES) undermined the performance of visual localization in deaf cats, while in hearing cats the deactivation only undermined the performance of auditory localization but not visual localization (M. A. Meredith et al., 2011). These findings are strong evidence that deprivation of hearing can rewire the auditory cortex for processing visual instead of auditory inputs.

In the next section, I will focus on the interaction between auditory and visual processing in the auditory cortex of animal models. Compared to other neurophysiological techniques, extracellular recording of spikes and local field potentials (LFPs) provides both the finest temporal resolution and the highest spatial specificity, making it the most ideal for studying cross-modal interaction at neuronal level.

1.4. Interplay of auditory and visual inputs in cortical auditory neurons

1.4.1. Visual modulation of neuronal activities in the auditory cortex

Two pioneer studies have investigated neuronal activities associated with visual stimuli in the core and the belt areas of macaque auditory cortex. The first study adopted audiovisual conspecific vocalization as stimuli (Ghazanfar et al., 2005) and found that auditory evoked local field potentials (LFPs) were enhanced by co-presenting the audio of vocalizations and the corresponding video of facial movements. More than 70% of sampled core area neurons and 80% of belt area neurons were found visually modulated. When using a disk that mimics the mouth opening, the visual enhancement was weaker but still detectable, suggesting that low-level visual processing is sufficient in modulating responses to vocalization. In the second study, 1200-ms audiovisual movies, consisting of both vocalization and artificial stimuli, were used to identify visual modulations of auditory evoked spikes and LFPs (Kayser et al., 2008). The results showed LFP measurements were more sensitive to visual modulation than spike measurement. It was also shown that visual suppression of firing rates lasted almost twice as long as visual enhancement on post-stimulus time histogram, suggesting that excitatory and inhibitory visual inputs to the primary auditory cortex may involve different neural circuits.

Several studies have demonstrated visually-evoked spike activities in the auditory cortex using small animal models (ferret: Bizley et al., 2007; mouse: Hunt et al., 2006; ferret: M Alex Meredith & Allman, 2015; cat: M Alex Meredith & Lomber, 2011; cat: Merrikhi et al., 2022). One study investigated visually evoked spike activities in six sub-areas of ferret auditory cortex (Bizley et al., 2007). Although visually responsive neurons were discovered in all six areas, the proportions of visually responsive neurons in two core auditory areas (the primary auditory cortex, A1, and

the anterior auditory field, AAF) were the lowest. The presence of visually responsive neurons in ferret A1 and AAF was also documented in another study (M Alex Meredith & Allman, 2015).

In cat, visual responses in neurons located in the dorsal zones of auditory cortex (DZ) was examined using prolonged 500-ms noise stimulus (Merrikhi et al., 2022), and it was reported that ~75% of the examined neurons were affected by visual influence. Similarly, in the auditory field of anterior ectosylvian sulcus (fAES), about 25% of neurons were also responsive to visual stimuli (M Alex Meredith et al., 2011). Various studies have previously reported a lack of visually responsive neurons in cat A1 (Kral, Schroder, Klinke, & Engel, 2003; Rebillard, Carlier, & Pujol, 1977; Stewart & Starr, 1970), but a recent study discovered that about 47.5% of examined neurons in A1 were visually responsive (Merrikhi et al., 2023).

1.4.2. Hearing loss unmask visual responses

The activation of neurons in the auditory cortex by non-auditory stimuli became even more outstanding in deaf animals. One study compared visually-evoked spike activities in anterior auditory cortex (AAF) between normal hearing ferrets and ferrets with 30 dB hearing loss (M. A. Meredith, Keniston, & Allman, 2012). Although visually responsive neurons only accounted for less than 3% of the total examined neurons in hearing ferrets, this proportion was increased to more than 38% in hearing-impaired ferrets. It was also noticed that while most multisensory neurons in hearing ferrets were suppressed by non-auditory stimuli, there were almost equivalent numbers of neurons showing cross-modal enhancement and suppression in the hearing loss ferrets. Both spontaneous and stimulus evoked firing rates were higher in hearing loss ferrets than in normal hearing ferrets. Similarly, there was no visually responsive neurons in AAF of hearing cats,

while 47% of AAF neurons became responsive to visual stimuli in deaf cats (M Alex Meredith & Lomber, 2011). These neurons were also examined with moving dots of different directions and velocities and demonstrated direction preference and a selectivity to high velocity.

Cat cross-modal plasticity was not only observed in the core areas of auditory cortex but also other auditory cortical areas. It was shown that visual responses were found in 67.7% of neurons in fAES in deaf cats, which is much higher than hearing cats (~25%) (M Alex Meredith et al., 2011). These neurons have a receptive field majorly located in the contralateral visual field with a small part extended into the ipsilateral visual field. When fAES was deactivated using cryoloop, normal hearing cats can no longer perform auditory localization but can still perform visual localization tasks, while deaf cats cannot perform the same visual localization tasks. These findings altogether suggested a functional conversion of fAES from auditory space coding to visual space coding (M Alex Meredith et al., 2011).

1.4.3. Relation to cognitive functions

In human and NHP subjects, some studies have found association between audiovisual interactions in the auditory cortex and cognitive functions (Lakatos, Karmos, Mehta, Ulbert, & Schroeder, 2008; Noppeney, 2021; Senkowski, Schneider, Foxe, & Engel, 2008; Talsma, Senkowski, Soto-Faraco, & Woldorff, 2010). It has been shown that visually evoked gamma-band oscillations of local field potentials (LFPs) in supragranular layers of the primary auditory cortex was enhanced by attending to visual stimuli while ignoring auditory stimuli, in comparison to attending to auditory stimuli while ignoring visual stimuli (Lakatos et al., 2008). This finding

suggested that audiovisual interaction in the auditory cortex involves top-down modulation of inter-cortical networks between the auditory system and non-auditory system.

The effect of aging on audiovisual interaction has been reported (Diederich, Colonius, & Schomburg, 2008; Stephen, Knoefel, Adair, Hart, & Aine, 2010). A human MEG study demonstrated less visual enhancement of auditory responses in superior temporal gyrus (STG), where the association auditory cortex resides, in the elder participants compared to young participants (Stephen et al., 2010). Behaviorally, the race model inequality has been tested in the same study and the violation of the race model was only observed in the young participants but not in the elder participants (Stephen et al., 2010), suggesting audiovisual interaction declines with aging.

Examples of audiovisual interaction participating in cognitive processing is also revealed by a line of studies, where language material is used as material. Orthographic-phonological interactions (i.e., interaction between the writing and the pronunciation of a syllable) refers to the audiovisual integration of language-related materials, which often involves cortical processing (Xu, Kolozsvári, Oostenveld, Leppänen, & Hämäläinen, 2019). It was shown with fMRI that activation in the auditory cortex was enhanced by semantically congruent letter-sound pairs while attenuated by incongruent letter-sound pairs, and the authors proposed that this congruency effect is derived from feedback inputs from STS/STG (Van Atteveldt, Roebroeck, & Goebel, 2009).

1.5. The role of temporal disparity in audiovisual processing

1.5.1. Temporal disparity in multisensory processing

One of the most widely discussed research topics on multisensory processing is the temporal integration of multisensory inputs. Behaviorally, the perception of auditory and visual stimuli as simultaneous against non-simultaneous (i.e. simultaneity judgement) and the perception of one modality before the other (i.e. temporal order judgement) is mainly determined by the temporal disparity of the auditory and visual stimuli (R. L. J. van Eijk, A. Kohlrausch, J. F. Juola, & S. van de Par, 2008).

When temporal audiovisual disparity is introduced, additional information regarding the relative timing of the two stimuli is available. Therefore, studies with temporal disparity as an experiment factor can be further divided into two categories depending on whether subjects are explicitly requested to process the relative temporal information. I will first introduce the temporal processing of audiovisual delay (**1.5.2**) as a unique category of studies before further discuss how temporal disparity affects sensory processing (**1.5.3**).

1.5.2. Temporal processing with audiovisual temporal disparity

Primarily, the temporal processing of audiovisual disparity is investigated as the perceived relative timing between an auditory stimulus and a visual stimulus using psychophysical approaches. In a temporal order judgement (TOJ) task, subjects are asked to identify the stimulus in which of the two different modalities occurred first in time. In a simultaneity judgement (SJ) task, subjects are asked to judge whether stimuli in two different modalities occurred simultaneously or not. Either way, the relative timing perceived by subjects were discretized into

a binary response. The probability of a positive response (as opposed to a negative response) can be estimated as a function of temporal disparity (i.e., psychometric function), where several features can be extracted. Temporal binding window (TBW) of audiovisual integration is commonly characterized by perceived point of simultaneity and just noticeable difference. The point of subjective simultaneity (PSS) is the temporal disparity (or audiovisual stimulus onset asynchrony, SOA) for the highest proportion of the trials where subjects report audiovisual stimuli were simultaneous. It sometimes can also be defined as the SOA for an estimated 50% of trials where subjects report auditory (visual) stimulus came first (Vroomen & Keetels, 2010). The just noticeable difference (JND) can be determined as the range of SOA subjects reporting simultaneous audiovisual stimuli higher at a proportion higher than the half of that at PSS, or the range of SOA subjects reporting auditory (visual) first more than 25% but less than 75% of the trials.

Temporal binding windows (TBW) is affected by various factors, such as the type of task (SJ versus TOJ) and stimulus type (speech, tools, flash/beep) (Stevenson & Wallace, 2013; R. L. Van Eijk, A. Kohlrausch, J. F. Juola, & S. van de Par, 2008). Studies have shown that training can effectively improve the performances in SJ and/or TOJ tasks or changed the characteristics of audiovisual integration (Fujisaki, Shimojo, Kashino, & Nishida, 2004; Stevenson, Wilson, Powers, & Wallace, 2013). The temporal binding windows (TBWs) could be shifted by exposure to asynchronous audiovisual impulse trains (Fujisaki et al., 2004) or speech (Navarra et al., 2005; Van der Burg & Goodbourn, 2015; Vatakis, Navarra, Soto-Faraco, & Spence, 2007). The width of temporal binding windows could be narrowed through training, which seemed to involve both multisensory and unisensory cortices (Powers, Hevey, & Wallace, 2012). In addition, musicians performed better in detecting delayed audiovisual pairs as deviants from synchronized pairs than

non-musicians (Lu, Paraskevopoulos, Herholz, Kuchenbuch, & Pantev, 2014), which suggested that TBWs could be shaped by experiences. In support of the malleability of TBWs, an age effect on perceived simultaneity and binding window widthas well as their changes after rapid re-calibration was reported (Noel, De Niear, Van der Burg, & Wallace, 2016). Finally, it has been shown that the performance in these two tasks are associated with several diseases (Wallace & Stevenson, 2014). Particularly, it is noteworthy that there is evidence showing visual temporal processing training (with two flashes) could decrease audiovisual TBWs as well (Stevenson et al., 2013), which suggested that the temporal processing of multisensory and uni-sensory inputs may have shared underlying mechanism.

TJ and SOJ tasks are very similar to each other, but an fMRI study revealed that they recruit different neural networks (Miyazaki et al., 2016). Contrasting blood-oxygen-level-dependent (BOLD) activation in the same subjects between the two tasks showed that several areas (left frontal, parietal and temporo-occipital regions) were more activated during the TOJ task than the SJ task (Binder, 2015). It has also been shown that performance in TOJ task declined more with aging in comparison with SJ task (Basharat, Adams, Staines, & Barnett-Cowan, 2018). It has been shown that rodents can perform these two tasks as well (Schormans et al., 2017).

Secondly, the temporal processing of audiovisual interaction was also identified in a temporal ventriloquism effect. Co-presenting auditory stimulus seems to cause an illusory auditory-going shift of the perceived time of visual stimulus. Stekelenburg and Vroomen (2005) using lag-effect threshold showed that a leading auditory stimulus was able to advance the perceived time of a lagging flash. In this paradigm, the perceived time of the lagging flash was quantified as the proportion of trials where a flash was perceived as lagged behind a bar horizontally moving across the location of the flash. Together with studies using TOJ and SJ tasks,

these behavioral results indicate that the perceived absolute and relative timing of auditory and visual stimuli are subject to multisensory interactions.

1.5.3. Non-temporal processing of audiovisual temporal disparity

The effect of cross-modal modulation also depends on audiovisual temporal disparity, regardless of whether or not participants are explicitly asked to make a judgement on the temporal order of audiovisual stimuli. The temporal factor of multisensory integration, ever since it was identified in SC neurons (M Alex Meredith et al., 1987), has been studied at both behavioral and physiological levels.

The behavioral benefit of audiovisual integration demonstrated by race model inequality has been extended on audiovisual stimuli with temporal disparity. A few studies showed that only physically simultaneous audiovisual stimuli (SOA=0) could induce the violation of the race model (Diederich & Colonius, 2004; Harrar, Harris, & Spence, 2017; L. Leone & McCourt, 2012; L. M. Leone & McCourt, 2013). However, one study found that multisensory benefit in the performance of a visual search task tolerated audiovisual disparity up to 100-msec (Van der Burg, Olivers, Bronkhorst, & Theeuwes, 2008).

When race model inequality and temporal window of audiovisual integration were modeled together, it was shown that elder participants had a broader integration window and larger multisensory enhancement than young participants (Colonius & Diederich, 2004; Diederich et al., 2008). It was also observed that the effect of audiovisual congruency on the identification of multisensory objects (e.g., sound and mouth movement of speech made by speaker) required a on small temporal disparity (Massaro, Cohen, & Smeele, 1996).

The effect of audiovisual disparity has been studied in both human and animal models using electrophysiological approaches, with a variety of selections of audiovisual delays from tens to hundreds of milliseconds. In human participants, two studies using event-related potential (ERP) techniques investigated the effect of temporal disparity systematically varied below 100 milliseconds (Naue et al., 2011; Thorne, De Vos, Viola, & Debener, 2011). Thorne et al. (2011) found that inter-trial phase coherence occurred for optimal visual-leading disparity between 30 and 75 msec. Naue et al. (2011) revealed auditory suppression of visually evoked beta oscillations with an auditory-leading disparity greater than 55 msec. The authors also pointed out that the modulation pattern followed a periodic pattern for disparity between 40 and 70 msec at a step of 5 msec, and that this periodicity resided in the beta band of EEG, suggesting that the auditory modulation of visual activations is through the phase-reset of ongoing EEG oscillation.

In animal models, the effect of audiovisual temporal disparity has been investigated at neuronal levels. Early study in cat superior colliculus demonstrated that multisensory enhancement decreased or even switched to suppression as temporal disparity increased (M Alex Meredith et al., 1987). In macaques, it has been observed that neuronal responses of A1 neurons to an auditory impulse can be suppressed by a preceding visual impulse. The visual suppression was modulated by visual-to-auditory delay, and the optimal delay for visual suppression was found between 20 and 80 ms (Kayser et al., 2008). In rat lateral extrastriate visual (V2L) cortex, a comparable optimal delays but for visual enhancements were identified (Schormans et al., 2017). In neurons from cat dorsal zone of auditory cortex (DZ), an unpublished study showed that the optimal delay for visual suppression of auditory onset responses was between 40- and 50-msec (Kok, 2015).

More studies adopted audiovisual delays beyond 100 msec, as well as a reduced precision (enlarged steps) in order to keep the experiment under a reasonable duration. This approach is

particularly common in studies using neuroimaging techniques (Franciotti, Brancucci, Della Penna, Onofrj, & Tommasi, 2011; Liu, Jin, Wang, & Gong, 2011; Lu et al., 2014; Stekelenburg & Vroomen, 2005). For example, Stekelenburg and Vroomen (2005) showed that the N1 amplitude of VEPs was smaller (less negatively-going) for 100-ms-leading click and larger for 100-ms-lagging click compared to the no-lag control. Using fMRI, it was shown that multisensory superior temporal cortex revealed two distinct components, where the first component showed facilitated BOLD signal only for synchronous auditory and visual speech (i.e., speaker's facial movement) (SOA=0) and the second component showed increasing BOLD signal as the temporal disparity increases (SOA=300, SOA=400) (Stevenson et al., 2010; Stevenson et al., 2011).

Finally, several studies have investigated the spectral composition of the temporal dynamics of multisensory interaction characterized by varying audiovisual delays. It has been suggested that different frequency bands of EEG oscillations may have distinct functional roles in sensory processing. For example, it has been suggested that the delta-theta band oscillation (<8 Hz) is involved anticipation and attention while the beta band oscillation (13-30 Hz) is involved in the error of prediction (Arnal and Giraud, 2012).

At behavioral level, Diederich, Schomburg, and Colonius (2012) showed that auditory modulation of saccadic reaction time (SRT) to a visual stimulus was affected by auditory-to-visual delay of stimulus. In this study, temporal disparity of a pair of audiovisual stimuli was varied at steps of 2 msec. The spectral decomposition revealed that the modulation of SRT oscillated with the temporal disparity, following a trend that can be regarded as the summation of multiple periodic patterns. The frequencies of these periodic patterns were found to reside in beta, gamma, and high-theta/alpha bands of EEGs. Similar spectral decomposition anlaysis was also applied to simple reaction time data derived from a visual detection task, where auditory-to-visual delay was varied

between 0 and 6 seconds (Fiebelkorn et al., 2011). The results of this study showed that there was a periodic pattern below 1 Hz in auditory modulation of visual detection, which =could sustain beyond 5 seconds after the onset of auditory stimuli.

At physiological level, it has been proposed that auditory stimulus could modulate the excitability of visual cortex through phase-reset alpha oscillations (Romei, Gross, & Thut, 2012). In this study, it was shown that, when the delay from auditory stimulus to transcranial magnetic stimulation (TMS) was in-phase with a periodic pattern cycling at ~10 Hz, the TMS at occipital lobe was more likely to induce an illusory perception of flash, in comparison to the sound-to-TMS delays that were out of phase. It was therefore implied that the alpha-band EEG oscillation (8~12 Hz) played a mediating role. In support of this hypothesis, it was shown that the excitability of visual cortex could be modulated by ongoing alpha wave (Romei et al., 2008).

1.6. The present investigation

The present thesis focused on investigating how visual stimulus modulates the auditory responses in the auditory cortex of hearing cats and its dependency on audiovisual temporal disparity. Chapter 2 and 4 mainly focused on establishing methodologies, whereas Chapter 3 and 5 contained the major findings more directly related to the topic of audiovisual interaction.

In Chapter 2, a novel behavioral paradigm was developed to train awake, head-free cats to perform a fixation task to prevent the contamination of artifacts due to eye and head movements during electrophysiological recording. In addition, reliable visual stimulation depends on subject's proficiency in maintaining the fixation at a designated area on the screen, which should be ideally

made as small as possible. To this end, five cats were tested with a psychophysics task, where variable fixation period was appended with a weak visual cue for detection. The psychometric functions of stimulus duration were successfully obtained at three stimulus intensities. The results demonstrated that cats were capable of acquiring a task involving fixating and attentively watching a small fixation dot presented on an LCD monitor through positive reinforcement.

In chapter 4, auditory evoked potentials (AEPs) and visually evoked potentials (VEPs) derived from scalp-recorded EEG were recorded from five cats lightly anesthetized by dexmedetomidine. Isoflurane was administrated with up-stepping concentrations, which caused a incrementing suppression that was much greater in VEPs than AEPs. The observation of isoflurane blocking visual activities suggested that isoflurane or similar drugs are not suitable anesthetics for studying visual modulation of AEPs.

In Chapter 3 and Chapter 5, I investigated visual modulation of click elicited LFPs and AEPs, respectively, using a novel audiovisual stimulus paradigm. In Chapter 3, two cat subjects were chronically implanted with electrode matrices in the primary auditory cortex. Using a similar behavioral training routine as in Chapter 2, the subjects practised performing the same fixation task, and, at the same time, passively watching an occasionally flipping checkerboard at the background. To evoked auditory responses, a train of click pulses were played with stochastic intervals throughout each fixation period. The primary reason of using such click trains instead of single click pulse was to make each individual click less salient in the stimulus context. Extracellular recording was made as neural measurement investigating how checkerboard flip affects click-elicited local field potentials (LFPs).

In Chapter 5, thirteen cat subjects were lightly anesthetized by dexmedetomidine without using isoflurane, as isoflurane was shown to selectively suppress visual processing in Chapter 4. A similar stimulus paradigm as in Chapter 3 was used, with flipping of checkerboard replaced by flashing of an LED diode. EEG signals were obtained as neural measurement investigating how visual flash affects click-elicited auditory evoked potentials (AEPs).

Some previous studies used a pair of auditory and visual stimuli and showed an optimal visual-to-auditory delay below about 100-msec for the largest visual modulation of auditory activities (Bizley et al., 2007; Kayser et al., 2008; M Alex Meredith et al., 1987), which implies for the advantage of temporal proximity for multisensory integration. This pattern, however, was not found in some other multisensory studies (Lakatos et al., 2007; Schormans et al., 2017; Stevenson et al., 2010). Therefore, the aim of data analysis in Chapter 3 and 5 was also to characterize the association between visual-to-auditory delay and visual modulation of click evoked LFPs / AEPs.

In summary, it is for the first time that multisensory responses were measured with scalp-recorded EEGs and with extracellular recordings in awake cats. The audiovisual visual stimuli used in the present investigation differs from previous studies using paired auditory and visual stimuli. When using paired auditory and visual stimulus impulses, the increase of delay between the two pulses also drastically changes the stimulus context (i.e., one event for small delay and two events for large delay). The results of these investigations were expected to further our understanding on the cross-modal interaction in the auditory cortex.

General goal: To establish modern techniques in cat as an animal model of investigating audiovisual interaction

General hypothesis: Invasive and non-invasive electrophysiological recordings as represented by local field potentials (LFPs) and auditory-evoked potentials (AEPs) are modulated by visual stimuli in anesthetized and awake cats.

1.7. References

Alais, D., Newell, F., & Mamassian, P. (2010). Multisensory processing in review: from physiology to behaviour. *Seeing and perceiving, 23*(1), 3-38.

Alencar, C. D., Butler, B. E., & Lomber, S. G. (2019). What and how the deaf brain sees. *Journal of Cognitive Neuroscience, 31*(8), 1091-1109.

Allman, B. L., Keniston, L. P., & Meredith, M. A. (2009). Adult deafness induces somatosensory conversion of ferret auditory cortex. *Proceedings of the National Academy of Sciences, 106*(14), 5925-5930.

Alvarado, J. C., Stanford, T. R., Rowland, B. A., Vaughan, J. W., & Stein, B. E. (2009). Multisensory integration in the superior colliculus requires synergy among corticocollicular inputs. *Journal of Neuroscience, 29*(20), 6580-6592.

Basharat, A., Adams, M. S., Staines, W. R., & Barnett-Cowan, M. (2018). Simultaneity and temporal order judgments are coded differently and change with age: an event-related potential study. *Frontiers in Integrative Neuroscience, 12*, 15.

Bavelier, D., & Neville, H. J. (2002). Cross-modal plasticity: where and how? *Nature Reviews Neuroscience, 3*(6), 443-452.

Beauchamp, M. S., Nath, A. R., & Pasalar, S. (2010). fMRI-guided transcranial magnetic stimulation reveals that the superior temporal sulcus is a cortical locus of the McGurk effect. *Journal of Neuroscience, 30*(7), 2414-2417.

Bergan, J. F., & Knudsen, E. I. (2009). Visual modulation of auditory responses in the owl inferior colliculus. *Journal of Neurophysiology, 101*(6), 2924-2933.

Bhattacharjee, A., Amanda, J. Y., Lisak, J. A., Vargas, M. G., & Goldreich, D. (2010). Vibrotactile masking experiments reveal accelerated somatosensory processing in congenitally blind braille readers. *Journal of Neuroscience, 30*(43), 14288-14298.

Binder, M. (2015). Neural correlates of audiovisual temporal processing–comparison of temporal order and simultaneity judgments. *Neuroscience, 300*, 432-447.

Bizley, J. K., Nodal, F. R., Bajo, V. M., Nelken, I., & King, A. J. (2007). Physiological and anatomical evidence for multisensory interactions in auditory cortex. *Cerebral Cortex, 17*(9), 2172-2189.

Bronchti, G., Heil, P., Scheich, H., & Wollberg, Z. (1989). Auditory pathway and auditory activation of primary visual targets in the blind mole rat (Spalax ehrenbergi): I. 2-deoxyglucose study of subcortical centers. *Journal of Comparative Neurology, 284*(2), 253-274.

Butler, B. E., Chabot, N., & Lomber, S. G. (2016). Quantifying and comparing the pattern of thalamic and cortical projections to the posterior auditory field in hearing and deaf cats. *Journal of Comparative Neurology, 524*(15), 3042-3063.

Butler, B. E., de la Rua, A., Ward-Able, T., & Lomber, S. G. (2018). Cortical and thalamic connectivity to the second auditory cortex of the cat is resilient to the onset of deafness. *Brain Structure and Function, 223*, 819-835.

Cappe, C., Murray, M. M., Barone, P., & Rouiller, E. M. (2010). Multisensory facilitation of behavior in monkeys: effects of stimulus intensity. *Journal of Cognitive Neuroscience, 22*(12), 2850-2863.

Caron-Desrochers, L., Schönwiesner, M., Focke, K., & Lehmann, A. (2018). Assessing visual modulation along the human subcortical auditory pathway. *Neuroscience Letters, 685*, 12-17.

Chabot, N., Butler, B. E., & Lomber, S. G. (2015). Differential modification of cortical and thalamic projections to cat primary auditory cortex following early-and late-onset deafness. *Journal of Comparative Neurology, 523*(15), 2297-2320.

Cheng, L., Fei, X.-Y., & Qu, Y.-L. (2019). Visual input shapes the auditory frequency responses in the inferior colliculus of mouse. *Hearing Research, 381*, 107777.

Cheng, L., Guo, Z.-Y., & Qu, Y.-L. (2020). Cross-modality modulation of auditory midbrain processing of intensity information. *Hearing Research, 395*, 108042.

Cloke, J. M., Jacklin, D. L., & Winters, B. D. (2015). The neural bases of crossmodal object recognition in non-human primates and rodents: a review. *Behavioural Brain Research, 285*, 118-130.

Codina, C., Buckley, D., Port, M., & Pascalis, O. (2011). Deaf and hearing children: a comparison of peripheral vision development. *Developmental science, 14*(4), 725-737.

Colonius, H., & Diederich, A. (2004). Multisensory interaction in saccadic reaction time: a time-window-of-integration model. *Journal of Cognitive Neuroscience, 16*(6), 1000-1009.

Diederich, A., & Colonius, H. (2004). Bimodal and trimodal multisensory enhancement: effects of stimulus onset and intensity on reaction time. *Perception and Psychophysics, 66*(8), 1388-1404.

Diederich, A., Colonius, H., & Schomburg, A. (2008). Assessing age-related multisensory enhancement with the time-window-of-integration model. *Neuropsychologia, 46*(10), 2556-2562.

Diederich, A., Schomburg, A., & Colonius, H. (2012). Saccadic reaction times to audiovisual stimuli show effects of oscillatory phase reset. *PloS One, 7*(10), e44910.

Doron, N., & Wollberg, Z. (1994). Cross-modal neuroplasticity in the blind mole rat Spalax ehrenbergi: a WGA-HRP tracing study. *Neuroreport, 5*(18), 2697-2702.

Fetsch, C. R., Turner, A. H., DeAngelis, G. C., & Angelaki, D. E. (2009). Dynamic reweighting of visual and vestibular cues during self-motion perception. *Journal of Neuroscience, 29*(49), 15601-15612.

Fiebelkorn, I. C., Foxe, J. J., Butler, J. S., Mercier, M. R., Snyder, A. C., & Molholm, S. (2011). Ready, set, reset: stimulus-locked periodicity in behavioral performance demonstrates the consequences of cross-sensory phase reset. *Journal of Neuroscience, 31*(27), 9971-9981.

Finney, E. M., Clementz, B. A., Hickok, G., & Dobkins, K. R. (2003). Visual stimuli activate auditory cortex in deaf subjects: evidence from MEG. *Neuroreport, 14*(11), 1425-1427.

Finney, E. M., Fine, I., & Dobkins, K. R. (2001). Visual stimuli activate auditory cortex in the deaf. *Nature Neuroscience, 4*(12), 1171-1173.

Franciotti, R., Brancucci, A., Della Penna, S., Onofrj, M., & Tommasi, L. (2011). Neuromagnetic responses reveal the cortical timing of audiovisual synchrony. *Neuroscience, 193*, 182-192.

Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. y. (2004). Recalibration of audiovisual simultaneity. *Nature Neuroscience, 7*(7), 773-778.

Gao, C., Green, J. J., Yang, X., Oh, S., Kim, J., & Shinkareva, S. V. (2023). Audiovisual integration in the human brain: a coordinate-based meta-analysis. *Cerebral Cortex, 33*(9), 5574-5584.

Ghazanfar, A. A., Maier, J. X., Hoffman, K. L., & Logothetis, N. K. (2005). Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *Journal of Neuroscience, 25*(20), 5004-5012.

Gingras, G., Rowland, B. A., & Stein, B. E. (2009). The differing impact of multisensory and unisensory integration on behavior. *Journal of Neuroscience, 29*(15), 4897-4902.

Glennon, E., Svirsky, M. A., & Froemke, R. C. (2020). Auditory cortical plasticity in cochlear implant users. *Current Opinion in Neurobiology, 60*, 108-114.

Groh, J. M., Trause, A. S., Underhill, A. M., Clark, K. R., & Inati, S. (2001). Eye position influences auditory responses in primate inferior colliculus. *Neuron, 29*(2), 509-518.

Gruters, K. G., & Groh, J. M. (2012). Sounds and beyond: multisensory and other non-auditory signals in the inferior colliculus. *Frontiers in neural circuits, 6*, 96.

Hammond-Kenny, A., Bajo, V. M., King, A. J., & Nodal, F. R. (2017). Behavioural benefits of multisensory processing in ferrets. *European Journal of Neuroscience, 45*(2), 278-289.

Harrar, V., Harris, L. R., & Spence, C. (2017). Multisensory integration is independent of perceived simultaneity. *Experimental Brain Research, 235*, 763-775.

Harrison, J., & Irving, R. (1966). Visual and Nonvisual Auditory Systems in Mammals: Anatomical evidence indicates two kinds of auditory pathways and suggests two kinds of hearing in mammals. *Science, 154*(3750), 738-743.

Hirokawa, J., Sadakane, O., Sakata, S., Bosch, M., Sakurai, Y., & Yamamori, T. (2011). Multisensory information facilitates reaction speed by enlarging activity difference between superior colliculus hemispheres in rats. *PloS One, 6*(9), e25283.

Hoefer, M., Tyll, S., Kanowski, M., Brosch, M., Schoenfeld, M. A., Heinze, H.-J., & Noesselt, T. (2013). Tactile stimulation and hemispheric asymmetries modulate auditory perception and neural responses in primary auditory cortex. *Neuroimage, 79*, 371-382.

Hunt, D., Yamoah, E., & Krubitzer, L. (2006). Multisensory plasticity in congenitally deaf mice: how are cortical areas functionally specified? *Neuroscience, 139*(4), 1507-1524.

Irving, R., & Harrison, J. (1967). The superior olivary complex and audition: a comparative study. *Journal of Comparative Neurology, 130*(1), 77-86.

Jennings, S. G. (2021). The role of the medial olivocochlear reflex in psychophysical masking and intensity resolution in humans: a review. *Journal of Neurophysiology, 125*(6), 2279-2308.

Kayser, C. (2010). The multisensory nature of unisensory cortices: a puzzle continued. *Neuron, 67*(2), 178-180.

Kayser, C., & Logothetis, N. K. (2007). Do early sensory cortices integrate cross-modal information? *Brain structure and function, 212*, 121-132.

Kayser, C., Petkov, C. I., Augath, M., & Logothetis, N. K. (2005). Integration of touch and sound in auditory cortex. *Neuron, 48*(2), 373-384.

Kayser, C., Petkov, C. I., Augath, M., & Logothetis, N. K. (2007). Functional imaging reveals visual modulation of specific fields in auditory cortex. *Journal of Neuroscience, 27*(8), 1824-1835.

Kayser, C., Petkov, C. I., & Logothetis, N. K. (2008). Visual modulation of neurons in auditory cortex. *Cerebral Cortex, 18*(7), 1560-1574.

Kayser, C., Petkov, C. I., & Logothetis, N. K. (2009). Multisensory interactions in primate auditory cortex: fMRI and electrophysiology. *Hearing Research, 258*(1-2), 80-88.

Kok, M. A. (2015). *Exploring the Structural and Functional Organization of the Dorsal Zone of Auditory Cortex in Hearing and Deafness*: The University of Western Ontario (Canada).

Kok, M. A., Chabot, N., & Lomber, S. G. (2014). Cross-modal reorganization of cortical afferents to dorsal auditory cortex following early-and late-onset deafness. *Journal of Comparative Neurology, 522*(3), 654-675.

Komura, Y., Tamura, R., Uwano, T., Nishijo, H., & Ono, T. (2005). Auditory thalamus integrates visual inputs into behavioral gains. *Nature Neuroscience, 8*(9), 1203-1209.

Kral, A., Schroder, J. H., Klinke, R., & Engel, A. K. (2003). Absence of cross-modal reorganization in the primary auditory cortex of congenitally deaf cats. *Experimental Brain Research, 153*(4), 605-613. doi:10.1007/s00221-003-1609-z

Lakatos, P., Chen, C.-M., O'Connell, M. N., Mills, A., & Schroeder, C. E. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron, 53*(2), 279-292.

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science, 320*(5872), 110-113.

Lazard, D. S., & Giraud, A.-L. (2017). Faster phonological processing and right occipito-temporal coupling in deaf adults signal poor cochlear implant outcome. *Nature Communications, 8*(1), 14872.

Lee, A. K., Wallace, M. T., Coffin, A. B., Popper, A. N., & Fay, R. R. (2019). *Multisensory Processes: The Auditory Perspective* (Vol. 68): Springer.

Lee, D. S., Lee, J. S., Oh, S. H., Kim, S.-K., Kim, J.-W., Chung, J.-K., . . . Kim, C. S. (2001). Cross-modal plasticity and cochlear implants. *Nature, 409*(6817), 149-150.

Leone, L., & McCourt, M. E. (2012). *The question of simultaneity in multisensory integration.* Paper presented at the Human Vision and Electronic Imaging XVII.

Leone, L. M., & McCourt, M. E. (2013). The roles of physical and physiological simultaneity in audiovisual multisensory facilitation. *i-Perception, 4*(4), 213-228.

Lessard, N., Paré, M., Lepore, F., & Lassonde, M. (1998). Early-blind human subjects localize sound sources better than sighted subjects. *Nature, 395*(6699), 278-280.

Levänen, S., & Hamdorf, D. (2001). Feeling vibrations: enhanced tactile sensitivity in congenitally deaf humans. *Neuroscience Letters, 301*(1), 75-77.

Liu, B., Jin, Z., Wang, Z., & Gong, C. (2011). The influence of temporal asynchrony on multisensory integration in the processing of asynchronous audio-visual stimuli of real-world events: an event-related potential study. *Neuroscience, 176*, 254-264.

Lomber, S. G., Meredith, M. A., & Kral, A. (2010). Cross-modal plasticity in specific auditory cortices underlies visual compensations in the deaf. *Nature Neuroscience, 13*(11), 1421-1427.

Lomber, S. G., Meredith, M. A., & Kral, A. (2010). Cross-modal plasticity in specific auditory cortices underlies visual compensations in the deaf. *Nature Neuroscience, 13*(11), 1421-U1163. doi:10.1038/nn.2653

Lu, Y., Paraskevopoulos, E., Herholz, S. C., Kuchenbuch, A., & Pantev, C. (2014). Temporal processing of audiovisual stimuli is enhanced in musicians: evidence from magnetoencephalography (MEG). *PloS One, 9*(3), e90686.

Macaluso, E., & Driver, J. (2005). Multisensory spatial interactions: a window onto functional integration in the human brain. *Trends in Neurosciences, 28*(5), 264-271.

Macaluso, E., George, N., Dolan, R., Spence, C., & Driver, J. (2004). Spatial and temporal factors during processing of audiovisual speech: a PET study. *Neuroimage, 21*(2), 725-732.

Martuzzi, R., Murray, M. M., Michel, C. M., Thiran, J.-P., Maeder, P. P., Clarke, S., & Meuli, R. A. (2007). Multisensory interactions within human primary cortices revealed by BOLD dynamics. *Cerebral Cortex, 17*(7), 1672-1679.

Massaro, D. W., Cohen, M. M., & Smeele, P. M. (1996). Perception of asynchronous and conflicting visual and auditory speech. *The Journal of the Acoustical Society of America, 100*(3), 1777-1786.

McDonald, J. J., Teder-Salejarvi, W. A., & Hillyard, S. A. (2000). Involuntary orienting to sound improves visual perception. *Nature, 407*(6806), 906-908. doi:Doi 10.1038/35038085

Meijer, G. T., Pie, J. L., Dolman, T. L., Pennartz, C. M., & Lansink, C. S. (2018). Audiovisual integration enhances stimulus detection performance in mice. *Frontiers in Behavioral Neuroscience, 12*, 231.

Meredith, M. A., & Allman, B. L. (2015). Single-unit analysis of somatosensory processing in the core auditory cortex of hearing ferrets. *European Journal of Neuroscience, 41*(5), 686-698.

Meredith, M. A., Keniston, L. P., & Allman, B. L. (2012). Multisensory Dysfunction Accompanies Crossmodal Plasticity Following Adult Hearing Impairment. *Neuroscience, 214*, 136-148. doi:10.1016/j.neuroscience.2012.04.001

Meredith, M. A., Kryklywy, J., McMillan, A. J., Malhotra, S., Lum-Tai, R., & Lomber, S. G. (2011). Crossmodal reorganization in the early deaf switches sensory, but not behavioral roles of auditory cortex. *Proceedings of the National Academy of Sciences of the United States of America, 108*(21), 8856-8861. doi:10.1073/pnas.1018519108

Meredith, M. A., Kryklywy, J., McMillan, A. J., Malhotra, S., Lum-Tai, R., & Lomber, S. G. (2011). Crossmodal reorganization in the early deaf switches sensory, but not behavioral

roles of auditory cortex. *Proceedings of the National Academy of Sciences, 108*(21), 8856-8861.

Meredith, M. A., & Lomber, S. G. (2011). Somatosensory and visual crossmodal plasticity in the anterior auditory field of early-deaf cats. *Hearing Research, 280*(1-2), 38-47.

Meredith, M. A., & Lomber, S. G. (2017). Species-dependent role of crossmodal connectivity among the primary sensory cortices. *Hearing Research, 343*, 83-91.

Meredith, M. A., Nemitz, J. W., & Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *Journal of Neuroscience, 7*(10), 3215-3229.

Merrikhi, Y., Kok, M. A., Carrasco, A., Meredith, M. A., & Lomber, S. G. (2022). Multisensory responses in a belt region of the dorsal auditory cortical pathway. *European Journal of Neuroscience, 55*(2), 589-610.

Merrikhi, Y., Kok, M. A., Lomber, S. G., & Meredith, M. A. (2023). A comparison of multisensory features of two auditory cortical areas: primary (A1) and higher-order dorsal zone (DZ). *Cerebral Cortex Communications, 4*(1), tgac049.

Middlebrooks, J. (2015). Auditory system: Central pathways.

Miyazaki, M., Kadota, H., Matsuzaki, K. S., Takeuchi, S., Sekiguchi, H., Aoyama, T., & Kochiyama, T. (2016). Dissociating the neural correlates of tactile temporal order and simultaneity judgements. *Scientific Reports, 6*(1), 23323.

Moro, S. S., Kelly, K. R., McKetton, L., Gallie, B. L., & Steeves, J. K. (2015). Evidence of multisensory plasticity: Asymmetrical medial geniculate body in people with one eye. *NeuroImage: Clinical, 9*, 513-518.

Murray, M. M., Thelen, A., Thut, G., Romei, V., Martuzzi, R., & Matusz, P. J. (2016). The multisensory function of the human primary visual cortex. *Neuropsychologia, 83*, 161-169.

Musacchia, G., Sams, M., Nicol, T., & Kraus, N. (2006). Seeing speech affects acoustic information processing in the human brainstem. *Experimental Brain Research, 168*, 1-10.

Musacchia, G., & Schroeder, C. E. (2009). Neuronal mechanisms, response dynamics and perceptual functions of multisensory interactions in auditory cortex. *Hearing Research, 258*(1-2), 72-79.

Nagy, A., Eördegh, G., Paróczy, Z., Márkus, Z., & Benedek, G. (2006). Multisensory integration in the basal ganglia. *European Journal of Neuroscience, 24*(3), 917-924.

Nath, A. R., & Beauchamp, M. S. (2012). A neural basis for interindividual differences in the McGurk effect, a multisensory speech illusion. *Neuroimage, 59*(1), 781-787.

Naue, N., Rach, S., Strüber, D., Huster, R. J., Zaehle, T., Körner, U., & Herrmann, C. S. (2011). Auditory event-related response in visual cortex modulates subsequent visual responses in humans. *Journal of Neuroscience, 31*(21), 7729-7736.

Navarra, J., Vatakis, A., Zampini, M., Soto-Faraco, S., Humphreys, W., & Spence, C. (2005). Exposure to asynchronous audiovisual speech extends the temporal window for audiovisual integration. *Cognitive Brain Research, 25*(2), 499-507.

Nishimura, H., Hashikawa, K., Doi, K., Iwaki, T., Watanabe, Y., Kusuoka, H., . . . Kubo, T. (1999). Sign language 'heard'in the auditory cortex. *Nature, 397*(6715), 116-116.

Noel, J.-P., De Niear, M., Van der Burg, E., & Wallace, M. T. (2016). Audiovisual simultaneity judgment and rapid recalibration throughout the lifespan. *PloS One, 11*(8), e0161698.

Noesselt, T., Tyll, S., Boehler, C. N., Budinger, E., Heinze, H.-J., & Driver, J. (2010). Sound-induced enhancement of low-intensity vision: multisensory influences on human sensory-specific cortices and thalamic bodies relate to perceptual enhancement of visual detection sensitivity. *Journal of Neuroscience, 30*(41), 13609-13623.

Noppeney, U. (2021). Perceptual inference, learning, and attention in a multisensory world. *Annual Review of Neuroscience, 44*, 449-473.

Nordmark, P. F., Pruszynski, J. A., & Johansson, R. S. (2012). BOLD responses to tactile stimuli in visual and auditory cortex depend on the frequency content of stimulation. *Journal of Cognitive Neuroscience, 24*(10), 2120-2134.

Powers, A. R., Hevey, M. A., & Wallace, M. T. (2012). Neural correlates of multisensory perceptual learning. *Journal of Neuroscience, 32*(18), 6263-6274.

Rauschecker, J., & Harris, L. (1983). Auditory compensation of the effects of visual deprivation in the cat's superior colliculus. *Experimental Brain Research, 50*, 69-83.

Rauschecker, J. P., & Kniepert, U. (1994). Auditory localization behaviour in visually deprived cats. *European Journal of Neuroscience, 6*(1), 149-160.

Rebillard, G., Carlier, E., & Pujol, R. (1977). Visual evoked responses on the primary auditory cortex in the cat after an early suppression of cochlear receptors (author's transl). *Revue D'electroencephalographie et de Neurophysiologie Clinique, 7*(3), 284-289.

Reig, R., & Silberberg, G. (2014). Multisensory integration in the mouse striatum. *Neuron, 83*(5), 1200-1212.

Röder, B., Rösler, F., & Neville, H. J. (1999). Effects of interstimulus interval on auditory event-related potentials in congenitally blind and normally sighted humans. *Neuroscience Letters, 264*(1-3), 53-56.

Romei, V., Brodbeck, V., Michel, C., Amedi, A., Pascual-Leone, A., & Thut, G. (2008). Spontaneous fluctuations in posterior α-band EEG activity reflect variability in excitability of human visual areas. *Cerebral Cortex, 18*(9), 2010-2018.

Romei, V., Gross, J., & Thut, G. (2012). Sounds reset rhythms of visual cortex and corresponding human visual perception. *Current Biology, 22*(9), 807-813.

Sandmann, P., Dillier, N., Eichele, T., Meyer, M., Kegel, A., Pascual-Marqui, R. D., . . . Debener, S. (2012). Visual activation of auditory cortex reflects maladaptive plasticity in cochlear implant users. *Brain, 135*(2), 555-568.

Schepers, I. M., Hipp, J. F., Schneider, T. R., Röder, B., & Engel, A. K. (2012). Functionally specific oscillatory activity correlates between visual and auditory cortex in the blind. *Brain, 135*(3), 922-934.

Schormans, A. L., Scott, K. E., Vo, A. M., Tyker, A., Typlt, M., Stolzberg, D., & Allman, B. L. (2017). Audiovisual temporal processing and synchrony perception in the rat. *Frontiers in Behavioral Neuroscience, 10*, 246.

Schürmann, M., Caetano, G., Hlushchuk, Y., Jousmäki, V., & Hari, R. (2006). Touch activates human auditory cortex. *Neuroimage, 30*(4), 1325-1331.

Scott, G. D., Karns, C. M., Dow, M. W., Stevens, C., & Neville, H. J. (2014). Enhanced peripheral visual processing in congenitally deaf humans is supported by multiple brain regions, including primary auditory cortex. *Frontiers in Human Neuroscience, 8*, 177.

Senkowski, D., Schneider, T. R., Foxe, J. J., & Engel, A. K. (2008). Crossmodal binding through neural coherence: implications for multisensory processing. *Trends in Neurosciences, 31*(8), 401-409.

Shim, H. J., Go, G., Lee, H., Choi, S. W., & Won, J. H. (2019). Influence of visual deprivation on auditory spectral resolution, temporal resolution, and speech perception. *Frontiers in Neuroscience, 13*, 1200.

Shore, S. E., & Zhou, J. (2006). Somatosensory influence on the cochlear nucleus and beyond. *Hearing Research, 216*, 90-99.

Spence, C., Senkowski, D., & Röder, B. (2009). Crossmodal processing. In (Vol. 198, pp. 107-111): Springer.

Stekelenburg, J. J., & Vroomen, J. (2005). An event-related potential investigation of the time-course of temporal ventriloquism. *Neuroreport, 16*(6), 641-644.

Stephen, J. M., Knoefel, J. E., Adair, J., Hart, B., & Aine, C. J. (2010). Aging-related changes in auditory and visual integration measured with MEG. *Neuroscience Letters, 484*(1), 76-80.

Stevens, C., & Neville, H. (2006). Neuroplasticity as a double-edged sword: Deaf enhancements and dyslexic deficits in motion processing. *Journal of Cognitive Neuroscience, 18*(5), 701-714.

Stevenson, R. A., Altieri, N. A., Kim, S., Pisoni, D. B., & James, T. W. (2010). Neural processing of asynchronous audiovisual speech perception. *Neuroimage, 49*(4), 3308-3318.

Stevenson, R. A., Kim, S., & James, T. W. (2009). An additive-factors design to disambiguate neuronal and areal convergence: measuring multisensory interactions between audio, visual, and haptic sensory streams using fMRI. *Experimental Brain Research, 198*, 183-194.

Stevenson, R. A., VanDerKlok, R. M., Pisoni, D. B., & James, T. W. (2011). Discrete neural substrates underlie complementary audiovisual speech integration processes. *Neuroimage, 55*(3), 1339-1345.

Stevenson, R. A., & Wallace, M. T. (2013). Multisensory temporal integration: task and stimulus dependencies. *Experimental Brain Research, 227*, 249-261.

Stevenson, R. A., Wilson, M. M., Powers, A. R., & Wallace, M. T. (2013). The effects of visual training on multisensory temporal processing. *Experimental Brain Research, 225*, 479-489.

Stewart, D. L., & Starr, A. (1970). Absence of visually influenced cells in auditory cortex of normal and congenitally deaf cats. *Experimental Neurology, 28*(3), 525-528.

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *The journal of the acoustical society of america, 26*(2), 212-215.

Szczepaniak, W. S., & Møller, A. R. (1993). Interaction between auditory and somatosensory systems: a study of evoked potentials in the inferior colliculus. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section, 88*(6), 508-515.

Taljaard, D. S., Olaithe, M., Brennan-Jones, C. G., Eikelboom, R. H., & Bucks, R. S. (2016). The relationship between hearing impairment and cognitive function: a meta-analysis in adults. *Clinical Otolaryngology*, *41*(6), 718-729. doi:10.1111/coa.12607

Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences, 14*(9), 400-410.

Thorne, J. D., De Vos, M., Viola, F. C., & Debener, S. (2011). Cross-modal phase reset predicts auditory task performance in humans. *Journal of Neuroscience, 31*(10), 3853-3861.

Van Atteveldt, N., Roebroeck, A., & Goebel, R. (2009). Interaction of speech and script in human auditory cortex: insights from neuro-imaging and effective connectivity. *Hearing Research, 258*(1-2), 152-164.

Van der Burg, E., & Goodbourn, P. T. (2015). Rapid, generalized adaptation to asynchronous audiovisual speech. *Proceedings of the Royal Society B: Biological Sciences, 282*(1804), 20143083.

Van der Burg, E., Olivers, C. N., Bronkhorst, A. W., & Theeuwes, J. (2008). Pip and pop: nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology: Human Perception and Performance, 34*(5), 1053.

Van Eijk, R. L., Kohlrausch, A., Juola, J. F., & van de Par, S. (2008). Audiovisual synchrony and temporal order judgments: Effects of experimental method and stimulus type. *Perception and Psychophysics, 70*, 955-968.

van Eijk, R. L. J., Kohlrausch, A., Juola, J. F., & van de Par, S. (2008). Audiovisual synchrony and temporal order judgments: Effects of experimental method and stimulus type. *Perception & Psychophysics, 70*(6), 955-968. doi:10.3758/Pp.70.6.955

Vatakis, A., Navarra, J., Soto-Faraco, S., & Spence, C. (2007). Temporal recalibration during asynchronous audiovisual speech perception. *Experimental Brain Research, 181*, 173-181.

Vos, T., Allen, C., Arora, M., Barber, R. M., Bhutta, Z. A., Brown, A., . . . Incidence, G. D. I. (2016). Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet, 388*(10053), 1545-1602. doi:DOI 10.1016/s0140-6736(16)31678-6

Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony: a tutorial review. *Attention, Perception, & Psychophysics, 72*(4), 871-884.

Wallace, M. T., & Stevenson, R. A. (2014). The construct of the multisensory temporal binding window and its dysregulation in developmental disabilities. *Neuropsychologia, 64*, 105-123.

Wong, C., Chabot, N., Kok, M. A., & Lomber, S. G. (2015). Amplified somatosensory and visual cortical projections to a core auditory area, the anterior auditory field, following early-and late-onset deafness. *Journal of Comparative Neurology, 523*(13), 1925-1947.

Xu, W., Kolozsvári, O. B., Oostenveld, R., Leppänen, P. H. T., & Hämäläinen, J. A. (2019). Audiovisual processing of Chinese characters elicits suppression and congruency effects in MEG. *Frontiers in Human Neuroscience, 13*, 18.

2. Chapter 2: A behavioral paradigm for extracellular recording in awake head-free cats

2.1. Relation to the thesis

There has been no literature documenting successful training of cat fixation. Here, we successfully trained multiple cat subjects in performing a fixation task using classical and operant conditionings with the assistance of a movement monitoring system customized with an inexpensive webcam.

This manuscript has been published on *Journal of Vision*.

Bao, X., Salloum, A., Gordon, S. G., & Lomber, S. G. (2020). The limited capacity of visual temporal integration in cats. *Journal of Vision*, 20(8), 28-28.

2.2. Abstract

It has been long known that prolonging stimulus duration may increase the perceived brightness of a visual stimulus. The interaction between intensity and duration generally follows a rule, such as that described in Bloch's law. This visual temporal integration relationship has been identified in human subjects and in non-human primates (NHPs). However, while auditory temporal integration has been extensively studied in the cat, visual temporal integration has not. Therefore, the goal of this study was to examine visual temporal integration in the cat. We trained five cats to respond when a brief luminance change was detected in a fixation dot. Following training, we measured the success rate of detecting the luminance change with varying durations

at threshold, subthreshold, and suprathreshold luminance levels. Psychometric functions showed that prolonging stimulus duration improved task performance, more noticeably for stimuli below 100 ms than beyond. Most psychometric functions were better fit to an exponential model than to a linear model. The gradually saturated performance observed here, as in previous studies, can be explained by the "leaky integrator" hypothesis, i.e. temporal integration is only valid below a critical duration. Overall, we developed a task whereby visual temporal integration was successfully demonstrated in the cat. The effect of stimulus duration on detection success rate displayed a pattern generally consistent with previous human and NHP findings on visual temporal integration.

2.3. Introduction

The temporal resolution of a sensory system is limited by the neurobiological basis of the receptor organ and the central nervous system. This can be reflected as a perceptual ambiguity between duration and intensity of a stimulus. For example, the human vision system, in some circumstances, interprets the prolongation of stimulus duration as an increase in brightness while stimulus luminance remains the same. Since it was initially documented 135 years ago (Bloch, 1885), this phenomenon has been extensively examined (Breitmeyer & Ganz, 1977; Gorea & Tyler, 1986; Kelly & Savoie, 1978; Rieiro et al., 2012; Roufs, 1974). In these studies, this phenomenon was often illustrated by iso-brightness curves in a stimulus luminance-duration domain, where a stimulus with long duration and low luminance has the same perceived brightness as another stimulus with short duration and high luminance. Such a trade-off between luminance

and duration shows that the visual system can integrate stimulus input over time to boost stimulus intensity at the expense of temporal resolution.

Meanwhile, as this trade-off is quantified with psychometric functions, many studies have identified that the trade-off ratio varies with stimulus duration. For example Aiba and Stevens (1964) found that the effect of doubling the stimulus duration on perceived brightness was equivalent to increasing luminance 10 times (decoupling stimulus luminance) for stimuli shorter than 10 ms, but became noticeably smaller for stimuli 10 to 100 ms long, and eventually could not be detected for stimuli longer than 100 ms. The marginal perceptual benefit of stimulus duration approaching zero, thus marking the "saturation" of the temporal integration window, is also known as the "critical duration". A few studies have reported a small decrease in the perceived brightness for medium-long duration stimuli (Gorea & Tyler, 1986; Kelly & Savoie, 1978; Rieiro et al., 2012). However, it is still being debated whether this phenomenon is due to a participant bias (Rieiro et al., 2012) or lateral inhibition (Gorea & Tyler, 1986, 2013).

A number of studies have examined the neural mechanism of visual temporal integration (Duysens, Gulyás, & Maes, 1991; Goris, Ziemba, Movshon, & Simoncelli, 2018; Levick & Zacks, 1970; Ohtani, Okamura, & Ejima, 2002; Scheich & Korn, 1971; White & Jeffreys, 1982). For example, it has been recently shown that an integrating neuron coding stimulus intensity by means of firing rate can be modulated by stimulus duration below a "neuronal" critical duration (Goris et al., 2018). Beyond that critical duration, such modulation became weaker and eventually absent.

Although these studies have shed some light on the neural correlates of temporal integration, future studies are still needed to explore its modality specificity and significance to behavior. To achieve these goals, it will be advantageous to have an animal model that is

convenient for both electrophysiological and behavioral studies in both the auditory and visual systems. It has already been shown that cat area 17 neurons demonstrate features of temporal integration (Duysens et al., 1991). However, to the best of our knowledge, no prior study has psychophysically examined visual temporal integration in the cat.

Overall, in this investigation we developed a visual task for the cat to establish a new animal model for studying temporal integration. The effect of stimulus duration on the success rate of detection was examined and displayed a pattern generally consistent with previous findings of visual temporal integration in humans and NHPs.

2.4. Methods

All procedures were conducted in compliance with the National Research Council's Guide for the Care and Use of Laboratory Animals (8th edition; 2011) and the Canadian Council on Animal Care's Guide to the Care and Use of Experimental Animals (1993), and adhere to the ARVO Animal Statement. Furthermore, the following procedures were also approved by the University of Western Ontario's Animal Use Subcommittee of the University Council on Animal Care.

Subjects

All subjects were domestic short-hair cats derived from a commercial breeding facility (Marshall BioResources, formerly Liberty Laboratories, Waverly, NY). The animals are group housed with other cats in an enriched environment with a 12-hour light cycle and ad libitum water. Their health was monitored by a veterinary technician on a daily basis and by a veterinarian once

per week. The facility was regularly inspected by the Ontario Ministry of Agriculture, Food, and Rural Affairs, and the Canadian Council on Animal Care.

Five cats (TRN, CHB, CTL, ARY, and BRN) were trained and tested in this experiment, and most of them are also participating in other projects. Concerning the other projects, TRN was surgically implanted with a multi-channel recording microelectrode in its left primary auditory cortex 2 months prior to testing for this experiment. ARY and BRN were deafened via daily subcutaneous injections of neomycin (60 mg/kg) for 30 days after birth (Leake, Hradek, Rebscher, & Snyder, 1991). No signs of weight loss or abnormal behavior except for the loss of hearing in ARY and BRN was noticed. Food was provided ad libitum for 1-hour per day after their daily training or testing.

Equipment and Software

Prior to training, the cats were conditioned to be loosely restrained by a canvas bag that was attached to a "cat chair", on which the cat could comfortably stand or sit while keeping their heads in a limited region facing towards a monitor. The monitor (XL2820, BenQ) was 59.8 cm wide by 33.6 cm high with a resolution of 1920 by 1080 and a refresh rate of 60 Hz and was placed horizontally ~40 cm in front of the cat. The visual stimulus was programmed and rendered in PsychToolbox on Matlab utilizing PCI graphic cards (Graphics Chipset AMD Radeon ® R9 200 Series).

Only positive reinforcement was used. Moist canned food was delivered with an auto-filled spoon placed between the cat and the monitor in such a way as to not block the cat's view of the monitor. At the time of reinforcement, a small amount of food was squirted into the spoon

through a hole at its bottom by a robotically pumped syringe under the electric control of a customized Arduino board that received digital commands from a computer.

The experiment was conducted in a dark acoustic chamber with infra-red illumination. Throughout the entire training or testing session, the cat was monitored by an infrared webcam by the experimenter outside the chamber. The camera data acquired was real-time analyzed using an adapted ROI-based motion detection algorithm using Computer Vision Toolbox (MathWorks) to identify specific cat behaviors that are crucial to the task, such as fixation and licking.

Stimulus and Procedures

A centered white dot on top of a constant gray background (0.3 cd/m2) was presented to the animal as a stimulus, as well as an orientation cue and a fixation cue, with the same diameter (32 pixel or 1.4 deg) but varying luminance. Stimulus luminance was calibrated with a photometer (Konica Minolta).

The orientation cue was a blinking white dot (50 cd/m2) switched on and off every 250 milliseconds. Once the animal was engaged with the fixation dot, a test trial was started by an experimenter pressing a button. On the monitor, the orientation cue became a fixation cue, which was the exact same white dot except for being constantly displayed instead of blinking. At any time within this fixation phase, if the movement of the cat's head in a designated camera area exceeded a pre-set threshold, the trial was restarted by return of the orientation cue.

After a pre-set temporal period (1250, 2000, or 2750 ms) of successful fixation, a test stimulus was presented by introducing a brief luminance change to the white fixation dot. The

duration and amplitude of the pulse of the luminance change were systematically manipulated as experiment variables, which will be referred to as pulse duration (PD) and pulse amplitude (PA).

After the onset of the testing stimulus, a 750 ms response window started to capture successful detection (i.e. cat un-fixating and approaching the spoon) or failed detection (i.e. cat either fixating or un-fixating without approaching the spoon). Following a successful detection (i.e., a hit), the spoon would be filled immediately as a reinforcement at the end of the trial. Following a failed detection (i.e., a miss), as long as the cat was still fixating (as in most cases), a second brief luminance change with the optimal PA and PD would be present, followed by a second 750 ms response window as well as the possible reinforcement, so as to strengthen the stimulus-response association and to maintain motivation (**Fig. 1**). False alarm can be found in some of the trials where subjects failed during the fixation period. In most of the trials, subjects had no problem passing the fixation period, which can be seen as correct rejection.

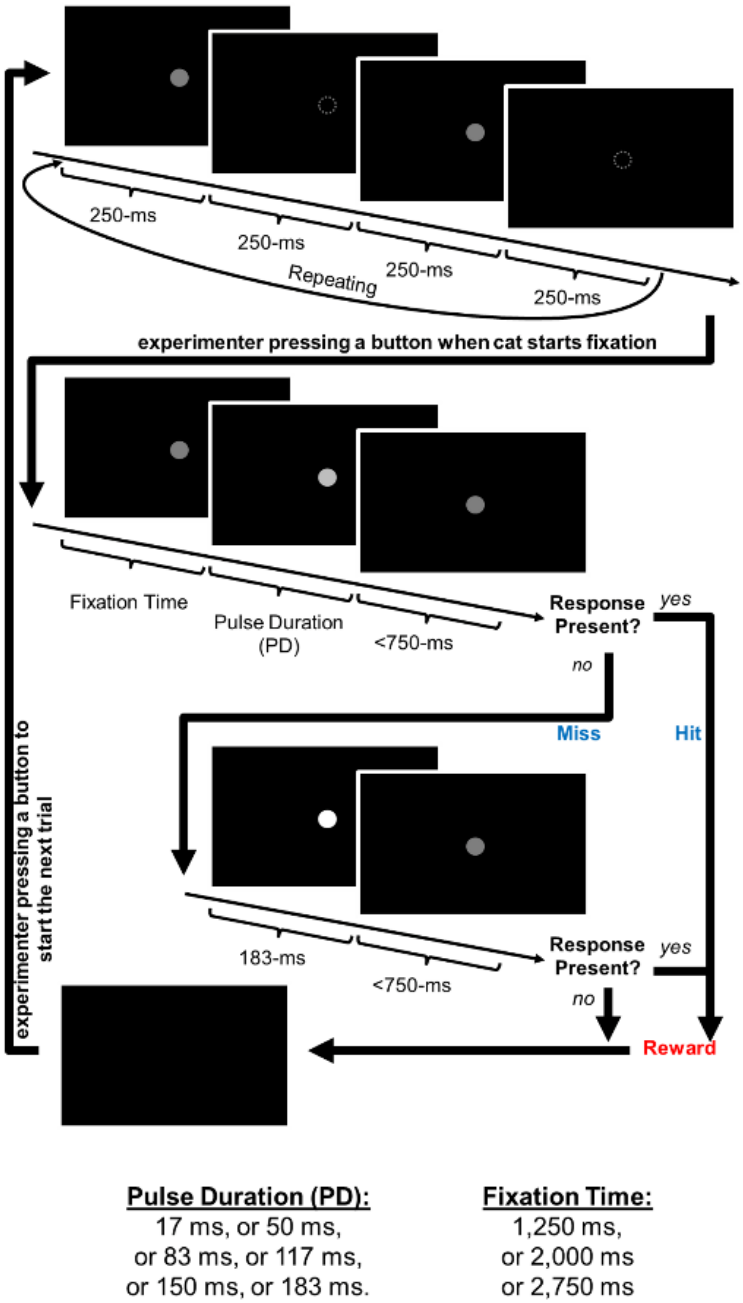


Figure 1. Schematic flow chart of a novel motivation-maintaining simple detection task.

Experimental Design

Cats were trained for 1 to 3 months before being able to perform the detection task at the optimal PA and PD with fixation time pre-set to 1250, 2000, 2750, or 3500 ms proficiently (success rate above 80%). Next, we determined a near-threshold PA for each cat by testing with varying PA between 0 and 70 cd/m2 (ref. to 50 cd/m2) while fixing the PD at 166 ms and looking at the psychometric function between success rate and PA.

Eventually, cats performed the detection task with PD randomly varying from 17 to 167 ms at the interval of 33 ms with the PA fixed at a threshold, subthreshold, or supra-threshold level. It was expected that the overall rates of detection would be lower with sub-threshold PA than near-threshold PA, but the effect PD may still be observed. A test block consisted of 18 trials where each of 6 PDs were tested for 3 cycles with the same PA. The order of the tested PDs was randomized for each cycle. Each trial was designated with one of three fixation times (1250, 2000, or 2750 ms), in a way that each PD was paired with each fixation time once and was used twice for one cycle. A typical testing session consisted of 3 testing blocks with 3 different PAs. The order of the tested PAs was counterbalanced across sessions. Each cat, except CTL, performed 6 to 18 testing sessions in total over 3 to 6 days.

Data Analysis

At the population level, we used a one-way repeated measure ANOVA to determine the statistical significance of the difference between the means of success rate of detection for different stimulus durations. This was conducted within the Statistics and Machine Learning Toolbox (MathWorks) in MATLAB. At an individual level, we also applied non-linear regressions to model the relationship between stimulus duration and success rate of detection using Curve Fitting

Toolbox (MathWorks). Parameters for different template functions were approximated using non-linear least squares method. The Levenberg-Marquardt algorithm was used during the fitting procedure. Goodness of fitting was converted to a z-score by Fisher transformation, and compared between linear and non-linear exponential models using a Student's t-test. The current data analysis did not include the trials when subjects failed to keep the fixation or made a response before stimulus onset, as the test stimulus was not present in these trials.

2.5. Results

Ability to detect a brief luminance deviant improved with pulse duration

In total, we obtained thirteen psychometric functions from five cats with sub-threshold, near-threshold, or supra-threshold PAs. On the averaged psychometric function, we found that (1) the rate of successful detection (i.e. success rate) increased monontonically with pulse duration (i.e. PD) and (2) the slope of this psychometric function decreased with PD (**Fig. 2**).

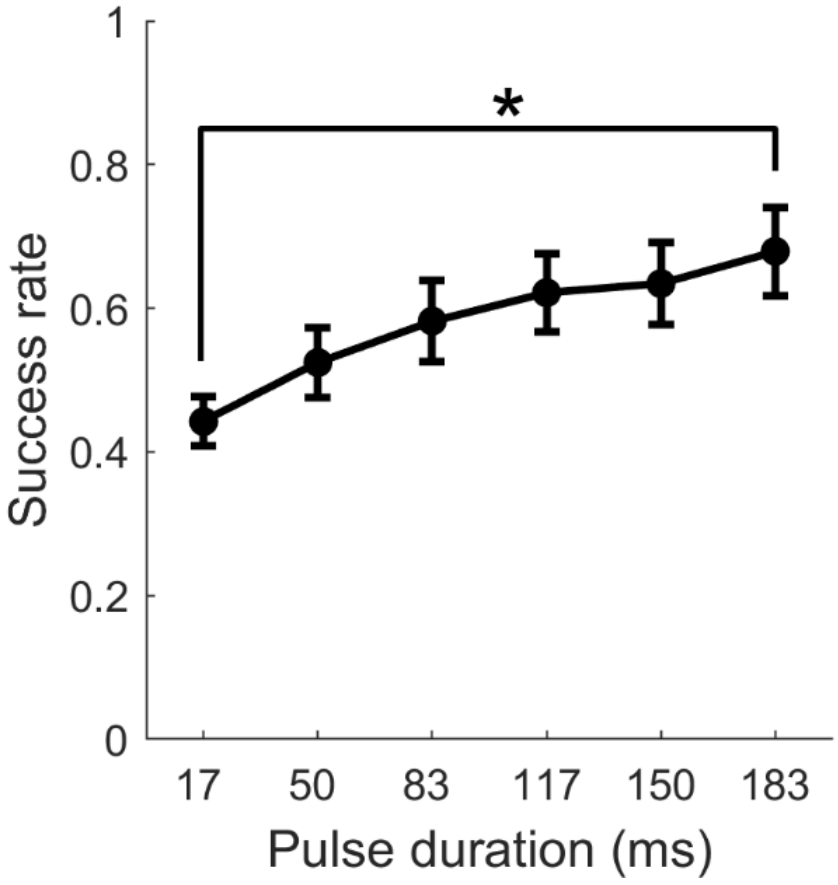


Figure 2. Grand average of success rate as a function of pulse duration. Success rate of detecting pulses of luminance change with sub-, near-, or supra-threshold levels of pulse amplitudes measured from 5 cats were averaged and plotted as a function of pulse duration (n = 13). Error bar, standard error (SE). *, significant difference shown by ANOVA post-hoc comparisons.

A one-way ANOVA showed that the main effect of PD was statistically significant ($F_{5, 72} = 2.61$, $p = 0.032 < 0.05$). This suggests that temporal integration does contribute to the detection of a weak visual stimulus. Next, we conducted post-hoc comparisons for all the possible combinations of two PDs. Significant difference was only observed between the shortest (17-ms) and the longest (183-ms) PD ($p = 0.027 < 0.05$).

To evaluate the effect of stimulus duration on the capacity of temporal integration, we calculated the increase in success rate of a 66-ms increase in stimulus duration, from 17 ms to 83 ms and from 117 ms to 183 ms, respectively. On average, the increase in success rate was larger for the short stimuli (~14%) than the long stimuli (~6%). Statistically, a paired two-sample Student's t-test showed that the difference between the two increases in success rate was marginally significant ($p = 0.058 > 0.05$).Altogether, our data indicated that the visual system in cats can integrate stimulus input over time to help in the detection of a weak visual signal that is behaviorally relevant. However, as observed in the cat auditory system (Costalupes, 1983; Solecki & Gerken, 1990), cat visual temporal integration has a limited temporal capacity to benefit detection."

Success rate saturated at long pulse durations

To further characterize such limited temporal capacity, we individually fitted each of the thirteen psychometric functions with both a linear and a non-linear customized model using fit function on MATLAB. The linear model used can be formulated as the following equation:

$$SR = a \cdot PD + b$$

where PD represents for all pulse durations used in the experiment design and SR represents for success rate respectively. Both *a* and *b* are free parameters and have no fitting bounds during the search of optimal fitness between the predicted and measured success rate. The non-linear exponential model used can be formulated as the following equation:

$$SR = a \cdot e^{-b \cdot PD} + c$$

where SR and PD again represent success rate and pulse duration acquired from the experiment, and e represents the base of the natural logarithm. To guide parameter searching, the upper bound

of parameter *a* was set to 0. The bottom and upper bounds of parameter c were set to 0 and 1. The curves generated by the non-linear model equation showed above using the best-fitting parameters for each of 13 testing sessions were plot and overlaid with experimentally measured success rates (**Fig. 3**).

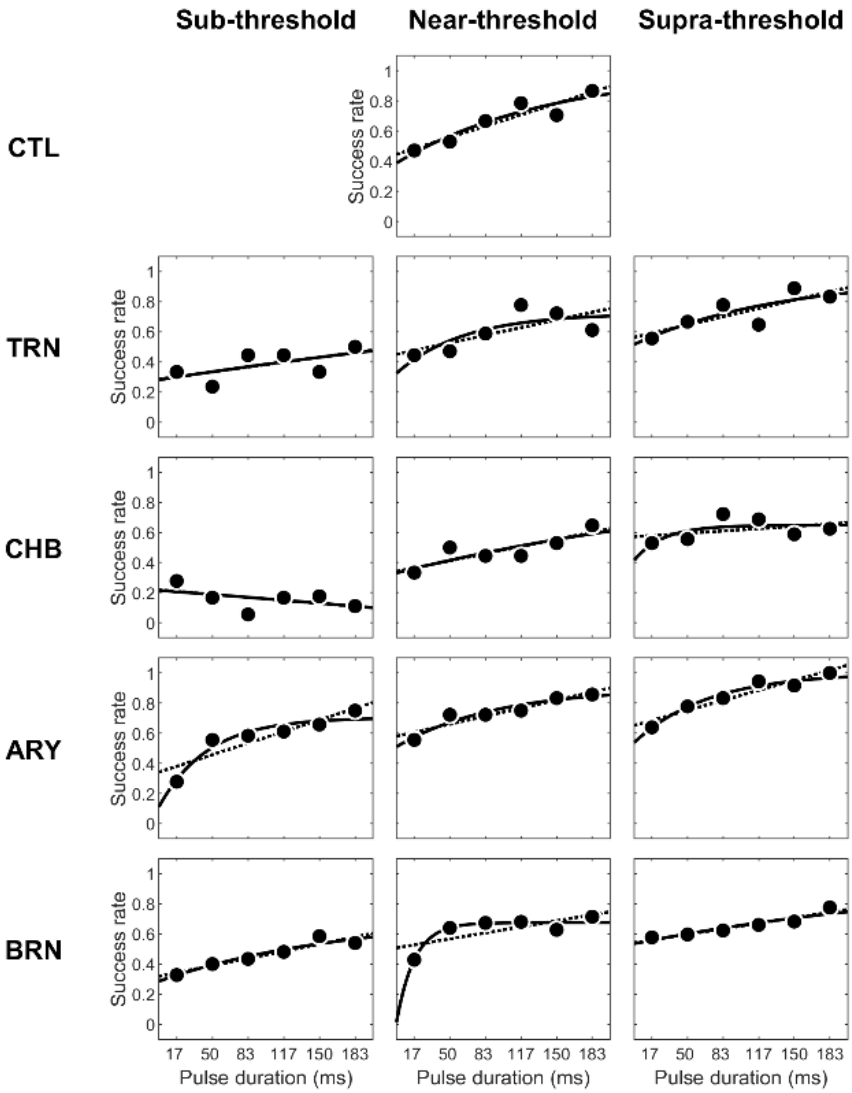


Figure 3. Individual psychometric functions. 13 Individual psychometric functions, with each of five subjects shown in different rows and each of three pulse amplitudes shown in different columns. In each plot, success rate was shown as a function of pulse duration, with experimental data indicated by filled dots and linear or non-linear best-fitting curve indicated by dashed or solid line.

Using R^2 as an index of goodness of fitting (GOF), we found that nine out of thirteen psychometric functions were fit better with the non-linear exponential model than with the linear model (**Fig. 4**). Overall, Fisher z-transformation of R^2 for the non-linear exponential model was significantly higher than that for the linear model ($p = 0.036 < 0.05$).

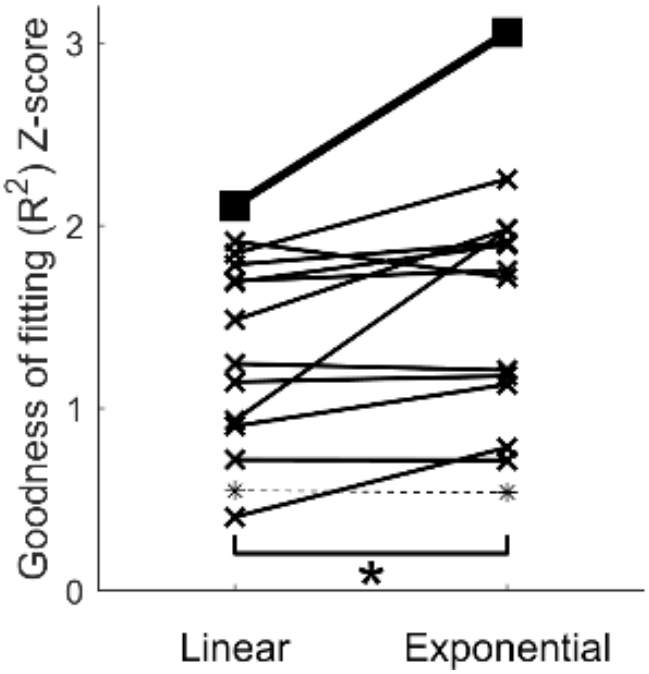


Figure 4. The comparison between linear and non-linear curve fitting. For each individual data fitting process (n = 12), except for the apparent outlier (*), goodness of fitting (r2) was converted into Fisher's Z-score and compared between the linear model and the non-linear exponential model, as indicated by (×). Statistical difference (*) between the two models were found (p < .05). The averaged psychometric function was also fit with both models (■).

These results suggest that the non-linear exponential model can be an authentic and useful simplification of the psychometric relationship between successful detection and stimulus duration. From the parameters derived from the fitting of the averaged data, it can be estimated

that the benefit of temporal integration for detecting a visual stimulus decays to 36.8% with each 101.4 ms increase of integration period.

As an attempt to explore the possibility of other non-linear models in fitting our experimental results, we also tried a sigmoid function and semi-logarithmic piecewise linear model as an alternative to an exponential function (**Fig. S1**). However, with our current dataset, these two models did not seem to demonstrate any benefit in goodness of fitting when compared with linear and exponential models (**Fig. S2**). It is also noted that these model have different numbers of free parameters and therefore the advantage of the exponential model (3 free parameters) over the linear model (2 free parameters) could have been an overestimate.

Saturated success rate continued to increase with pulse amplitude

To ensure that the saturation of the success rate was not a ceiling effect limited by cats' proficiency in performing the task in general, in four of the five cats, three different PAs were used. Therefore, we were also able to examine the effect of PA on success rate for each PD. Overall, three psychometric functions stacked up in the order of corresponding PA without crossing each other (**Fig. 5 Top**). A two-way (3×6) repeated measured ANOVA showed significant effects of both PD ($F_{5, 15} = 11.07$, $p < 0.001$) and PA ($F_{2, 6} = 30.74$, $p < 0.001$). Although no significant interaction was found, we applied the test of simple effect of PA at each level of PD (**Fig. 5 Bottom**). The effect of PA suggested that the task proficiency is not likely the reason for the saturated performance.

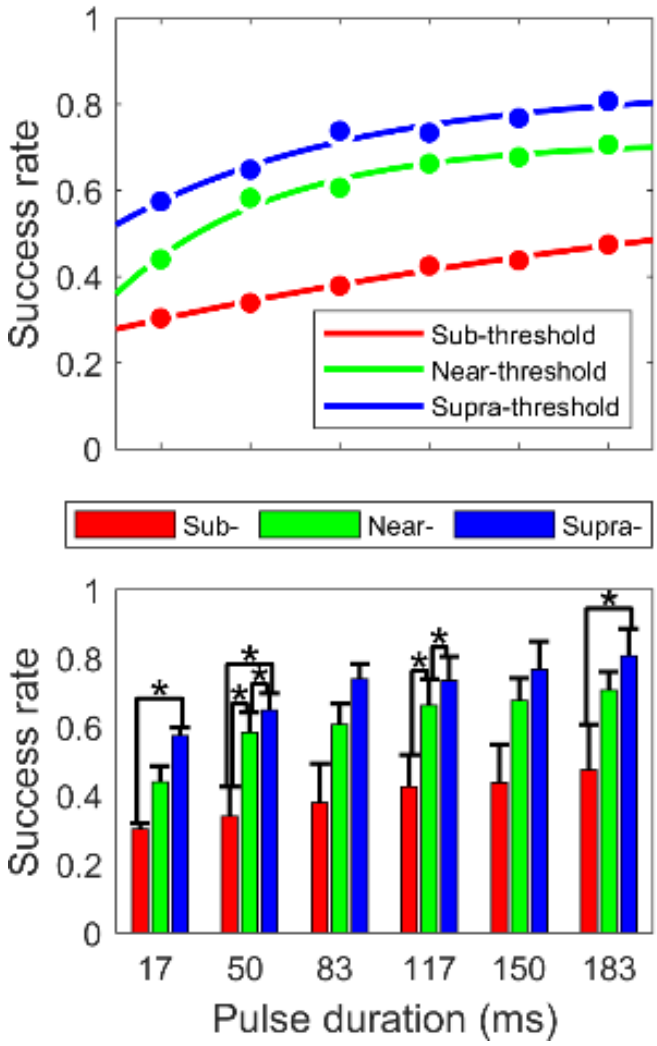


Figure 5. Effect of stimulus intensity on psychometric functions. Top, averaged psychometric functions for each of sub-, near-, supra-threshold pulse amplitudes were plot in red, green and blue, respectively. Bottom, bars with different colors (same as in top) compared averaged success rate for each of six pulse durations (n = 4). Errorbar, standard error (SE). *, p < .05 in ANOVA multiple comparison test.

We noticed that the trend we observed from population average was more consistently present for near- and supra-threshold than sub-threshold pulse amplitude (luminance) levels. Deaf subjects seemed to show more consistent individual psychometric functions than hearing subjects, although we can't make any conclusion from our limited number of subjects.

Summary

The data from this study showed that successful detection of a visual stimulus was modulated by the stimulus duration in the range between approximately 20 and 200 ms. At a population level, the effect of increasing stimulus duration on task performance was more prominent for short stimuli than long stimuli. For each individual testing session, the psychometric function was often better fit by a non-linear exponential model than a linear model. The duration effect and the unique shape of the psychometric functions persist with different stimulus intensities.

2.6. Discussions

Leaky integrator and energy integrator

Duysen (1991) used the terms “leaky integrator” and “energy integrator” to describe two of the most important aspects of temporal integration in area 17 neurons, which could also apply to behavioral measurements. A “leaky integrator” describes an input-output relationship where the input is added together over a finite-length time window. With intensity constant over time, prolonging the duration of the input (i.e. stimulus) increases the output (i.e. perception) before saturation is reached, which defines the critical duration. An “energy integrator”, on the other hand, takes both intensity and time as indirect factors of the output. Instead, the product of intensity and time, energy, is the direct factor. This term emphasizes the interchangeability of intensity and time in temporal integration (Duysens et al., 1991).

In human subjects, many previous studies have characterized visual temporal integration by measuring the intensity (luminance or contrast) thresholds as a function of stimulus duration (Gorea & Tyler, 1986; Rashbass, 1970) or the intensity level at each stimulus duration that is perceptually equivalent to a standard stimulus with both parameters fixed (Aiba & Stevens, 1964; Rieiro et al., 2012; Stévens & Hall, 1966). Both approaches allow for the observation of an "energy integrator" as an iso-brightness curve, where any stimulus featuring an intensity-duration pair on the curve is equally detectable or equally bright. While these studies have provided valuable insights in understanding visual temporal integration, there are some shortcomings as well. First, compared to estimating response rates on average, it takes many more trials to measure stimulus thresholds. Second, it is hard to avoid rise-decay slope co-varying with stimulus intensity when using a computer monitor to deliver the stimulus. Thus, given the goal of this study, it is sufficient, and probably more straight-forward, to use a stimulus with fixed intensity levels and measuring success rate of detection.

Bloch's law versus Broca-Sulzer's law

While almost all relevant studies reported a generally monotonic increase of perceived brightness with increasing stimulus duration as in Bloch's law (Bloch, 1885; Harwerth, Boltz, & Smith 3rd, 1980), a handful of studies have also reported a small negative marginal effect of increasing stimulus duration for medium-long duration stimuli (Breitmeyer & Ganz, 1977; Broca & Sulzer, 1902; Kelly & Savoie, 1978; Rieiro et al., 2012; Roufs, 1974). This finding was first documented and named in 1902 (Broca & Sulzer, 1902). Recently, the underlying mechanism for the difference observed in Broca-Sulzer's law compared to Bloch's law was debated (Gorea & Tyler, 2013). An earlier study proposed that Broca-Sulzer's law could be the result of inhibition, by showing that reducing the spatial frequency of a grating stimulus could effectively modify the

shape of the psychometric function observed in Bloch's law to that of Broca-Sulzer's law (Gorea & Tyler, 1986). The simulation in this study showed that the extra "dips" observed in Broca-Sulzer's law arises from a second inhibitory phase following excitation in the impulse-response function (the Fourier transform of a band-passing temporal modulation transfer function), which is a characteristic for grating stimuli of spatial frequency lower than 3-cyc/deg but not higher (Robson, 1966). However, a more recent study (Rieiro et al., 2012) proposed that subject bias is the reason for observing Bloch's law rather than Broca-Sulzer's law. They observed, in the same group of participants, Bloch's law with a blocked design and Broca-Sulzer's law with au unblocked design. Riero and colleagues (2012) therefore implied that Broca-Sulzer's law is a fundamental feature of visual system whereas Bloch's law demonstrates a perceptual representation of sensory input bias as in "brightness constancy" (Gilchrist et al., 1999).

In the present experimental design, the same luminance paired with different stimulus durations was used for each testing session. This design was similar to the blocked design used in an earlier experiment where Broca-Sulzer's law was observed (Rieiro et al., 2012). However, the shape of the averaged psychometric function from our data set was closer to Bloch's law. As individual psychometric functions tend to be noisy, it is difficult to determine if they favor either law. Also, we don't know if subject bias observed in human subjects would be comparable in the cat visual system. In a future study it would be interesting to compare our results from cats with those from human subjects.

Neural correlations of temporal integration in sensory systems

It is also important to note that such intensity-duration trade-off is not unique to the visual system. For instance, Stévens and Hall (1966) showed that both visual brightness and auditory

loudness grew with stimulus duration and eventually reached saturation (Stévens & Hall, 1966). Similarly, for tactile perception, overestimation of speed of whisker vibration occuring by simply increasing stimulus duration has been identified in both rats (Fassihi, Akrami, Pulecchi, Schönfelder, & Diamond, 2017) and monkeys (Luna, Hernández, Brody, & Romo, 2005). Fassihi and colleagues (2017) showed that neuronal activities after stimulus offset in vM1 (vibrissal motor cortex) but not vS1 (vibrissal sensory cortex) is modulated by both the speed and duration of tactile vibration. In this study, tactile stimuli with longer durations led to higher speed of vibration perceived by rats, and higher firing rates recorded from vM1 neurons. This finding suggests that temporal integration of sensory input may also be processed outside of primary sensory cortices. However, in this particular case, the activity of vM1 neurons may be highly task-specific or involved in the motor planning in response to sensory experience, rather than the sensory experience itself. Although the temporal integration of sensory input is shown to be largely accounted for by the integrative properties of sensory receptors, as well as sensory neurons in each individual sensory modality, it has been speculated that cortical processing may play a role at longer timescales (Mongillo & Loewenstein, 2017).

Several potential neural mechanisms of visual temporal integration have been proposed based on electrophysiological investigations. Typically, in vivo extracellular recording has been performed along the visual pathway to record neuronal firing in response to visual stimuli of varying duration. Examining the maximum neuronal firing rate, Duysens and colleagues (1991) showed that neuronal activity in cat area 17 were modulated by stimulus durations up to 80 ms, which is 50% ~ 100% higher than those previously reported in retinal ganglion cells (Levick & Zacks, 1970; Scheich & Korn, 1971). Duysens et al. (1991) was the first study to highlight the role of visual cortex in temporal integration, where the time scale is very close to that found in

behavioral studies. Another way of investigating temporal integration is to present a fixed-length stimulus (e.g. 500-ms) but iteratively analyze neural activities in a varying-length window aligned to the stimulus onset, if we can assume the effect of stimulus after the end of response window is negligible (but see Enns and Di Lollo (2000)). Neurometric functions constructed using this method for neurons in monkey primary visual cortex (V1) and medial temporal area (MT) showed that the marginal benefit of prolonging the response window on neuron's capacity of orientation discrimination decreased over time (Goris et al., 2018).

Auditory and visual temporal integration

Auditory stimuli, such as pure tones and noise, have been used in studies of temporal integration in many animal models (see Heil (2017) for a review), including cats (Costalupes, 1983; Solecki & Gerken, 1990). In cats, detection thresholds of sound levels for pure tones were measured for stimulus durations ranging from 50 to 1000 milliseconds. The sound level thresholds were found to decay with stimulus duration; with decay constants ranging between 100 and 1000 ms that were inversely correlated with pure tone frequency (Costalupes, 1983). Unfortunately, no comparable study investigating visual temporal integration using the same animal model could be identified. One of the obstacles that discourages such a study is the difficulty of engaging a cat in fixation that is sufficiently stable over the entire period of stimulus delivery.

In this study, we implemented a training paradigm that can quantify animal behavior with online image processing, and thus successfully trained five cats to perform a simple fixation-detection task. In this task, the cats were only loosely restrained in a canvas bag with a large freedom of movement in their head positions. Taking advantage of this training paradigm, we were able to reliably measure a cat's ability to detect brief pulses of luminance change in the fixation

dot. By constructing the psychometric functions of varying pulse duration, we were able to quantify the capacity for visual temporal integration in cat. For future studies, this task can be easily adapted into a bi-modal version, in which case a direct comparison between auditory and visual temporal integration can be made in the same subject.

Conclusion

In summary, considering the behavioral and electrophysiological findings in the current and previous studies, we believe that the cat can be a promising model for future research to answer unresolved issues in visual temporal integration.

2.7. References

Aiba, T. S., & Stevens, S. S. (1964). Relation of brightness to duration and luminance under light- and dark-adaptation. *Vision Research,* 4(7-8), 391-401. doi:10.1016/0042-6989(64)90011-2

Bloch, A.-M. (1885). Experiences sur la vision. Essai d'Optique sur la gradation de la lumie're Comptes Rendus de Séances de La Société de Biologie, Paris, 37, 493-495.

Breitmeyer, B. G., & Ganz, L. (1977). Temporal studies with flashed gratings: Inferences about human transient and sustained channels. *Vision Research,* 17(7), 861-865. doi:10.1016/0042-6989(77)90130-4

Broca, A., & Sulzer, D. (1902). La sensation lumineuse en fonetion du temps. *Journal de Physiologie et de Pathologie Generale,* 4, 632-640.

Costalupes, J. A. (1983). Temporal integration of pure tones in the cat. *Hearing Research,* 9(1), 43-54. doi:10.1016/0378-5955(83)90133-8

Duysens, J., Gulyás, B., & Maes, H. (1991). Temporal integration in cat visual cortex: A test of bloch's law. *Vision Research,* 31(9), 1517-1528. doi:10.1016/0042-6989(91)90129-s

Enns, J. T., & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences,* 4(9), 345-352.

Fassihi, A., Akrami, A., Pulecchi, F., Schönfelder, V., & Diamond, M. E. (2017). Transformation of perception from sensory to motor cortex. *Current Biology,* 27(11), 1585-1596.e1586. doi:10.1016/j.cub.2017.05.011

Gilchrist, A., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X., . . . Economou, E. (1999). An anchoring theory of lightness perception. *Psychological Review,* 106(4), 795. doi:10.1037/0033-295x.106.4.795

Gorea, A., & Tyler, C. W. (1986). New look at Bloch's law for contrast. *Journal of the Optical Society of America A,* 3(1), 52-61. doi:10.1364/josaa.3.000052

Gorea, A., & Tyler, C. W. (2013). Dips and bumps: On bloch's law and the broca-sulzer phenomenon. *Proceedings of the National Academy of Sciences of the USA,* 110(15), E1330-E1330. doi:10.1073/pnas.1221807110

Goris, R. L. T., Ziemba, C. M., Movshon, J. A., & Simoncelli, E. P. (2018). Slow gain fluctuations limit benefits of temporal integration in visual cortex. *Journal of Vision,* 18(8), 8. doi:10.1167/18.8.8

Harwerth, R. S., Boltz, R. L., & Smith 3rd, E. L. (1980). Psychophysical evidence for sustained and transient channels in the monkey visual system. *Vision Research,* 20(1), 15-22. doi:10.1016/0042-6989(80)90137-6

Heil, P., Matysiak, A., & Neubauer, H. (2017). A probabilistic Poisson-based model accounts for an extensive set of absolute auditory threshold measurements. *Hearing Research,* 353, 135-161. doi:10.1016/j.heares.2017.06.011

Kelly, D. H., & Savoie, R. E. (1978). Theory of flicker and transient responses. III. An essential nonlinearity. *Journal of the Optical Society of America,* 68(11), 1481-1490. doi:10.1364/josa.68.001481

Leake, P. A., Hradek, G. T., Rebscher, S. J., & Snyder, R. L. (1991). Chronic intracochlear electrical stimulation induces selective survival of spiral ganglion neurons in neonatally deafened cats. *Hearing Research,* 54(2), 251-271. doi:10.1016/0378-5955(91)90120-x

Levick, W. R., & Zacks, J. L. (1970). Responses of cat retinal ganglion cells to brief flashes of light. *The Journal of Physiology,* 206(3), 677-700. doi:10.1113/jphysiol.1970.sp009037

Luna, R., Hernández, A., Brody, C. D., & Romo, R. (2005). Neural codes for perceptual discrimination in primary somatosensory cortex. *Nature Neuroscience,* 8(9), 1210-1219. doi:10.1038/nn1513

Mongillo, G., & Loewenstein, Y. (2017). Neuroscience: Formation of a percept in the rat cortex. *Current Biology,* 27(11), R423-R425. doi:10.1016/j.cub.2017.04.019

Ohtani, Y., Okamura, S., & Ejima, Y. (2002). Temporal summation of magnetic response to chromatic stimulus in the human visual cortex. *NeuroReport,* 13(13), 1641-1644. doi:10.1097/00001756-200209160-00014

Rashbass, C. (1970). The visibility of transient changes of luminance. *Journal of Physiology,* 210(1), 165-186. doi:10.1113/jphysiol.1970.sp009202

Rieiro, H., Martinez-Conde, S., Danielson, A. P., Pardo-Vazquez, J. L., Srivastava, N., & Macknik, S. L. (2012). Optimizing the temporal dynamics of light to human perception. *Proceedings of the National Academy of Sciences of the USA,* 109(48), 19828-19833. doi:10.1073/pnas.1213170109

Robson, J. G. (1966). Spatial and temporal contrast-sensitivity functions of the visual system. *Journal of the Optical Society of America,* 56(8), 1141-1142.

Roufs, J. A. J. (1974). Dynamic properties of vision—IV: Thresholds of decremental flashes, incremental flashes and doublets in relation to flicker fusion. *Vision Research,* 14(9), 831-851. doi:10.1016/0042-6989(74)90148-5

Scheich, H., & Korn, A. (1971). Timing properties and temporal summation in the retina. *Pflügers Archiv: European Journal of Physiology,* 327(1), 16-36. doi:10.1007/bf00634096

Solecki, J. M., & Gerken, G. M. (1990). Auditory temporal integration in the normal-hearing and hearing-impaired cat. *Journal of the Acoustical Society of America,* 88(2), 779-785. doi:10.1121/1.399727

Stévens, J. C., & Hall, J. W. (1966). Brightness and loudness as functions of stimulus duration. *Perception & Psychophysics*, 1(9), 319-327. doi:10.3758/bf03215796

White, M. J. M., & Jeffreys, D. A. (1982). Pattern-evoked potentials and Bloch's law. *Vision Research*, 22(8), 897-903. doi:10.1016/0042-6989(82)90026-8

2.8. Supplementary Information

Method of fixation training - Customization

The first step we took was to introduce canned food with a designated spoon to our cat subjects, as this allows the experimenters to efficiently reward wanted behavior and distract them from disturbance during further training. A canvas bag that can be attached to the metal pillars around the cat chair was used in our paradigm for loosely restraining the cat subjects to the designated positions without causing any discomfort. Once cats were used to being easily put on the canvas bag and associate this process with reward, they were transported to the acoustic chamber, where the beforementioned spoon, bottom-filled with canned food, was fixed in-between cat chair and stimulation monitor. The amount of canned food decreased as the cat became more used to the environment inside the chamber, until the cat appeared comfortable and calmly expecting canned food even when no canned food is given, which marks the completion of customization.

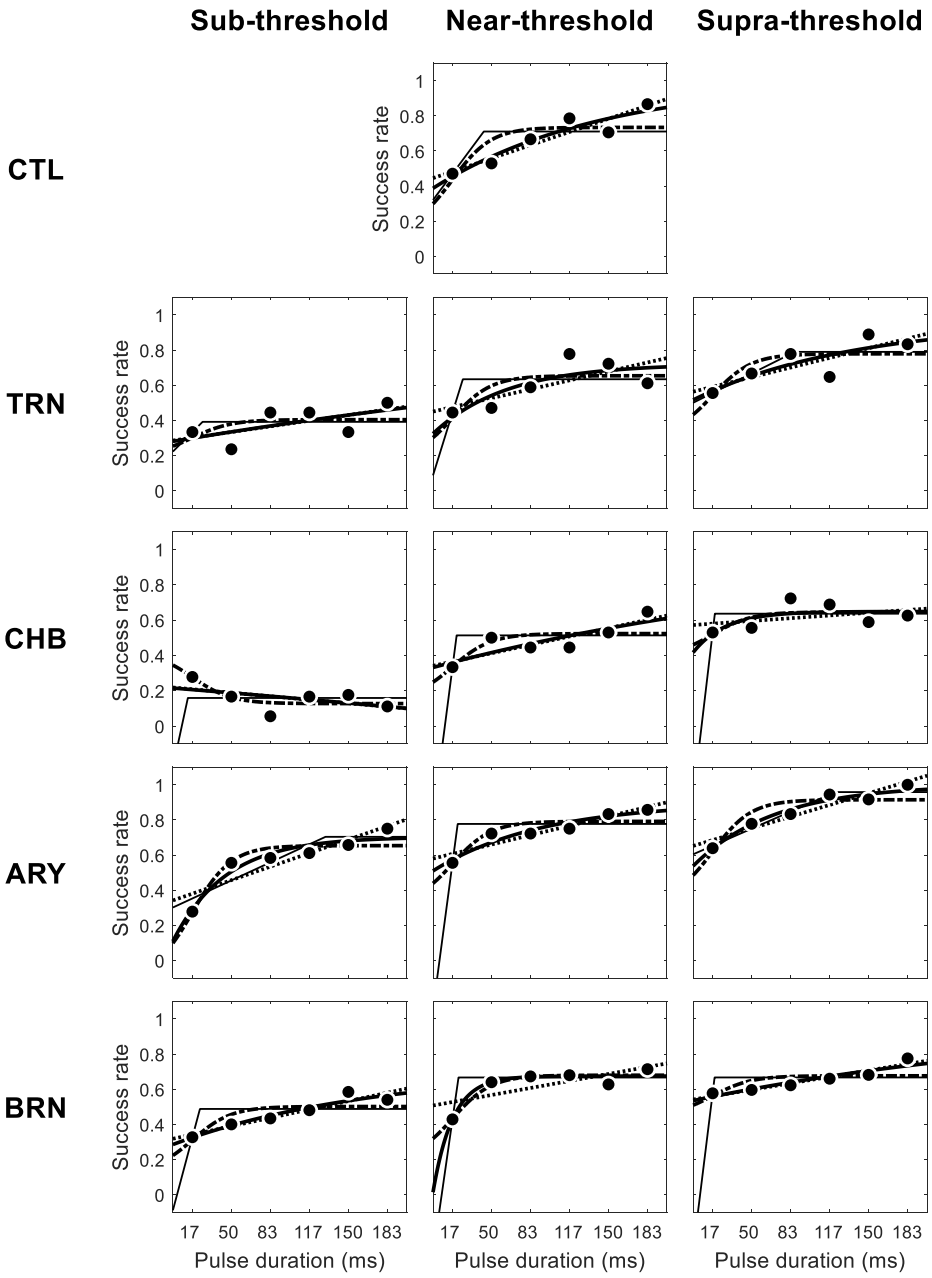
Method of fixation training - Association

In this phase, classic conditioning was used to associate a brighter fixation dot with canned food. A flashing fixation dot (at 2 Hz) was present to initiate a new trial. Once a cat starts fixating, the fixation dot will stop flashing and remain lit. While the cat is performing fixation, a digital trigger is sent to dispense a small amount of canned food and simultaneously present a brief (500-ms) pulse of increase in brightness of the fixation dot. Very quickly the pulse becomes a conditioned stimulus (CS) and the association is completed.

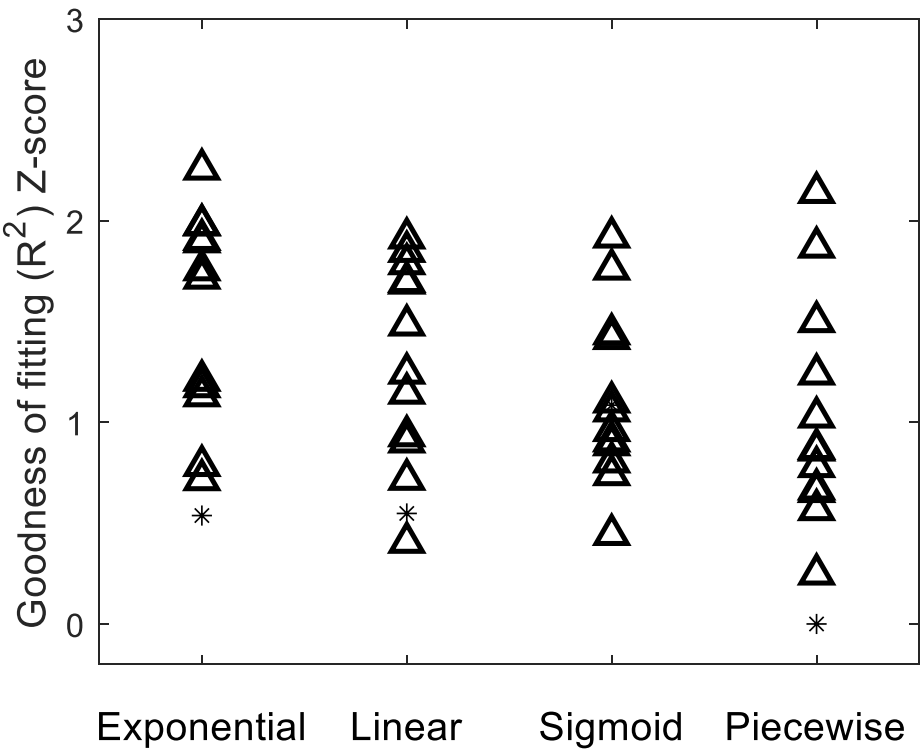
Method of fixation training - Refinement

In this phase, the task difficulty is increased whereas the fixation dot remains lit longer and longer before CS and reward. In a failed fixation trial (i.e., cat stops fixating before CS), the fixation dot starts flashing immediately as in a new trial. Any unwanted facial movement (e.g., eyes winking, mouth opening, tongue licking etc.) are considered a failure in fixation. In the framework of operant conditioning, the wanted behavior, which is fixating while fixation dot remain lit, is encouraged by CS and food, while unwanted behaviors can be discouraged by withholding reward.

As the proficiency of the fixation tasks improves in cats, monitoring of fixation performance is taken over by a real-time camera-assisted algorithm. Briefly, a region-of-interest (ROI) from the view of a web-came containing the cat's face is extracted and processed in Matlab for each frame to estimate the optical flow, which produces a real-time quantification of the cat's movement. Although the unit of this measurement is arbitrary and can be affected by the selection of ROI and lighting condition, its baseline during fixation and corresponding threshold it crosses during the response to CS can be quickly established and doesn't need much adjustment for days or even weeks in row.



Supplementary Figure 1. Individual psychometric functions fitting with linear and multiple non-linear models. As a supplementary result to Figure 1, three different non-linear models (thick solid lines, exponential function; dashdotted lines, sigmoid function; a piecewise function, thin solid lines) were used to fit success rate observed from our psychophysics experiment, as well the linear mode (dotted lines). To improve the stability in the fitting results, each fitting was repeated for 10 times and the one resulting the highest goodness of fitting was present here and kept for further analysis.



Supplementary Figure 2. The comparison of goodness of fitting between linear and multiple non-linear models. As a supplementary result to Figure 4, goodness of fitting (r^2) derived from three different non-linear models as well as the linear model was converted into Fisher's Z-score and present here. One-way ANOVA showed no significant main effect of modeling method on goodness of fitting.

3. Chapter 3: Visual modulation of neuronal responses to auditory inputs in awake cat primary auditory cortex

3.1. Relation to the thesis

Visual responses and visual modulation of auditory responses in the primary auditory cortex (A1) have been investigated in both non-human primates (NHPs) and small animal models (e.g., cats and ferrets) using extracellular recording. An adequate proportion of A1 neurons in NHPs demonstrated visual responses and an even larger proportion for visual modulation of auditory responses. Unlike regions outside core areas of cat auditory cortex, it has been reported in cat A1 that there is very little visual modulation and no visual response. A major methodological caveat between the two classes of studies lies on the anesthesia state of animal subjects. Visual responses documented in NHPs were observed in awake animal subjects, while the cat studies observing no visual response were performed when animal subjects were anesthetized.

In this chapter, it is for the first time that neuronal responses to audiovisual stimuli were recorded in cat A1 while subjects were not anesthetized. Two cats were behaviorally trained using similar methods as reported in Chapter 2, before being chronically implanted with electrode matrices in their left A1. The secondary part of novelty in this project is that a new audiovisual stimulus paradigm was designed for this experiment to characterize the role played by audiovisual temporal disparity (or, audiovisual delay), using the measurement of click-evoked local field potentials (LFPs). In this new paradigm, each individual acoustic stimulus (i.e., a click) was "blended" in the acoustic context of a stochastic click train, and therefore has less saliency compared to the traditional audiovisual stimuli that can be easily perceived as one or two events. The remove of acoustic saliency is a critical methodological consideration specifically for this

project, as awake subjects could be easily distracted from the fixation task by a sensory event, or, alternatively, would have to recruit additional cognitive process to ignore it.

The finding of this chapter supported the argument that visual stimulus can affect auditory responses in cat A1 neurons. The data also showed that the visual modulation in cat A1 depends on the delay of audiovisual stimulus. However, the pattern of delay dependency seemed rather individualized for different neurons. The optimal delays for visual modulation in each neuron often indicated multiple rather than single integration window. Brain oscillation driven by visual stimulus may contribute to the occurrence of multiple integration windows, although a wide-spread spectral distribution of brain oscillation rather than a specific EEG band may be involved.

The data of this manuscript has been partially presented on *International Multisensory Research Forum (IMRF).* The manuscript will be submitted with revision to *European Journal of Neuroscience*.

Bao, X., Lomber, S.G. (2023). Temporal Course of Visual Modulation Revealed by Local Field Potentials in Cat Primary Auditory Cortex. *International Multisensory Research Forum (IMRF).* (Podium #2 Talk 6)

3.2. Abstract

Understanding visual modulation of auditory processing in hearing individuals is an important step to unveil the mechanism of the cross-modal plasticity in deaf individuals and to

discover effective intervention for preserving multisensory functions after auditory restoration. It has been shown in various species that neurons in the primary auditory cortex (A1) also receive visual modulation, but there is still a lack of systematic investigation on how audiovisual temporal disparity affects the audiovisual interactions in A1. To follow up with this research question, we chronically implanted two 32-channel electrode matrices for extracellular recording from the primary auditory cortex (A1) in two cat subjects. During recording, the subjects performed a fixation task while being presented with audiovisual stimuli. Using a novel stimulus paradigm, our results showed that the visual modulation of click-evoked LFPs was affected by the delay between checkerboard flipping and click onset, with multiple temporal integration windows characterized from the same electrode. In another word, we observed multiple ranges instead of a single range of audiovisual delays where visual modulation could occur. While the attenuation of LFP power occurred most frequently for near-simultaneous audiovisual stimuli, the enhancement of LFP power was observed across a wide spanning of temporal disparities. The periodic pattern of the visual modulation was further analyzed, revealing a 60-Hz entrainment effect and delta band frequency coupling. Our findings are encouraging for future studies on characterizing visual modulation of deaf auditory cortex using cat as an animal model.

3.3. Introduction

The cross-modal interaction has been discovered in cat auditory cortex for more than one decade (Lomber, Meredith, & Kral, 2010). Retrograde tracers deposit in cat auditory cortices revealed possible inward projections from several visual areas (Butler, Chabot, & Lomber, 2016; Kok, Chabot, & Lomber, 2014; Meredith, Clemo, Corley, Chabot, & Lomber, 2016; Wong,

Chabot, Kok, & Lomber, 2015). Auditory neurons that also respond to visual stimulus were identified with extracellular recordings (Meredith et al., 2011; Meredith & Lomber, 2011; Merrikhi, Kok, Carrasco, Meredith, & Lomber, 2022). Moreover, both the anatomic connection and the functional characteristics suggested that the visual modulation in auditory cortices could be amplified by the experience of hearing loss, which implies that the multisensory domain of auditory cortex may play a critical role in the cross-modal plasticity following deafness (Finney, Clementz, Hickok, & Dobkins, 2003; Finney, Fine, & Dobkins, 2001; Lomber et al., 2010).

The observation of multisensory interaction specific to certain stimulus timing of two different modalities was laid down as one of the very first-determined guiding principles of multisensory processing (Meredith, Nemitz, & Stein, 1987). Ever since it was first discovered in cat superior colliculus (SC), the effect of stimulus timing was also demonstrated using extracellular recordings for audio-visual interaction in macaque core/belt auditory cortex (Kayser, Petkov, & Logothetis, 2008), ferret auditory cortex (Bizley, Nodal, Bajo, Nelken, & King, 2007), and rodent association auditory cortex (DZ) (Schormans et al., 2017), and for audio-tactile interaction in macaque primary auditory cortex (Lakatos, Chen, O'Connell, Mills, & Schroeder, 2007). Those studies have indicated that both cortical and subcortical areas carry the temporal information of multisensory stimuli for temporal processing that may underlie the perception of multisensory integration (Dixon & Spitz, 1980; Powers, Hillock, & Wallace, 2009; Stevenson & Wallace, 2013; Stevenson, Zemtsov, & Wallace, 2012; Wallace & Stevenson, 2014).

Meanwhile, evidence from brain imaging studies led to the change of notion that primary auditory cortices is solely for unisensory processing (Calvert et al., 1997; Kayser, Petkov, & Logothetis, 2009; Pekkola et al., 2005). Studies using extracellular recordings from the primary auditory cortex in animal models, in addition to those mentioned, are also in support that there

exists multisensory processing in a subset of neurons from the primary auditory cortex across different species, including macaques (Ghazanfar, Maier, Hoffman, & Logothetis, 2005; Kayser et al., 2008; Lakatos et al., 2007), ferrets (Bizley et al., 2007; Meredith & Allman, 2015) and rodents (Hunt, Yamoah, & Krubitzer, 2006; Kobayasi & Riquimaroux, 2013). Although the exact proportion varies from one study to another, in general local field potentials (LFPs) were more sensitive in detecting multisensory responses in the primary auditory cortices. Nevertheless, our previous knowledge about multisensory neurons in cat primary auditory cortex (A1) (Kral, Schröder, Klinke, & Engel, 2003; Rebillard, Carlier, Rebillard, & Pujol, 1977; Stewart & Starr, 1970) wasn't updated yet. The effect of stimulus timing in these neurons was also unknown as well.

The current study focused on cat A1 by means of click-evoked local field potentials (LFPs) under visual modulation with chronically implanted micro-electrode matrices. To characterize the effect of stimulus timing in awake subjects, we took an exploration on a novel stimulus paradigm that examines the effect of varying temporal audiovisual disparity, without change of stimulus salience during passive viewing. Here, we reported our findings from two subjects with four implanted electrode matrices and 128 electrodes in total. The visual modulation of click LFPs present in about a quarter of electrodes without any audiovisual correspondence. This proportion could have been improved by optimizing the delay between flip of checkerboard and click onset. The amount of visual modulation as a function of the delay featured complex temporal profile, as was shown in the delay function of audio-tactile interaction in macaque primary auditory cortex (Lakatos et al., 2007). Spectral decomposition served as a promising approach for explaining this complexity inside the framework laid by the "phase reset" hypothesis. In this framework, the excitability of sensory cortex is under the modulation of ongoing EEG oscillation, which can be

influenced by preceding sensory inputs from the same or different modality, as well as the mental pre-allocation such as selective attention (Calderone, Lakatos, Butler, & Castellanos, 2014; Lakatos et al., 2013).

.

3.4. Results

We made extracellular recording over multiple days from 64-channel electrode matrices in the left primary auditory cortex (A1) of two cat subjects trained to perform a fixation task. In each fixation trial, a train of clicks were played as auditory stimuli to drive the neuronal responses of A1. A checkerboard was displayed as the background of the fixation dot and flipped to its inverted version at random interval as visual stimulus. The visually evoked activities derived from averaging across trials were subtracted from single-trial signal (**Fig. 1**, see details in **Methods**). This stimulus and recording protocol allowed us to investigate the effect of visual modulation on click-evoked local field potentials (LFPs) and its dependency on audiovisual temporal disparity probed by flip-to-click delay. Unlike in a paired audiovisual stimulus paradigm, our novel stimulus paradigm was meant to prevent, to the largest extent, the confounding effect of stimulus saliency associated to varying delay between a visual pulse and an auditory pulse. This is done by playing multiple clicks as in a click train instead of just one click. As the entire trial is filled with clicks, it is unlikely that a certain unimodal or bimodal stimulus event can stand out among the others.

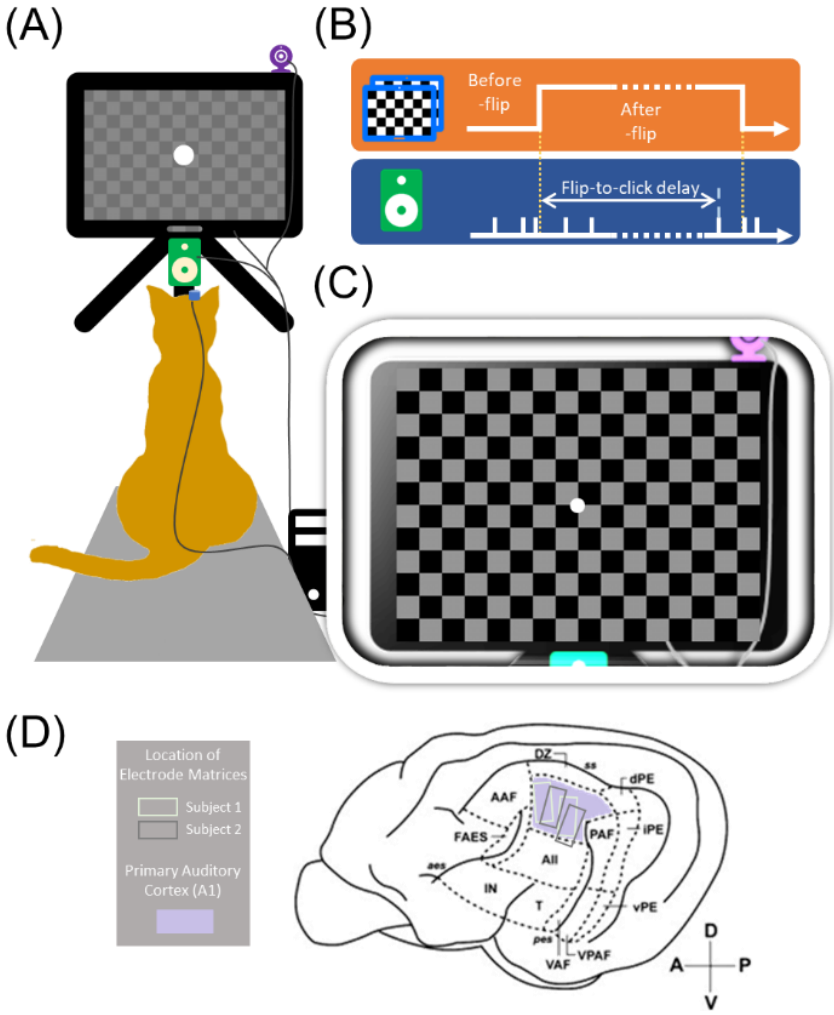


Figure 1. Experiment setup, stimulus design and electrode implantation. (A) The experiment setup for cat performing fixation task during extracellular recording. **(B)** The timeline of auditory and visual stimuli in an example trial. **(C)** A diagram showing locations of chronic implanted electrode matrices.

Local field potentials evoked by clicks before and after flip of checkerboard

As the timing of click pulses was configured in a stochastic manner using a Poisson random process, there was no systematic temporal correspondence between the auditory and visual stimuli. Therefore, an overall lack of cross-modal interactions was expected for the comparison between

click presented LFPs before and after flip of checkerboard. However, since the stimulus paradigm of this type was not tested at neuronal level by any previous studies, we still examined the effect of visual modulation regardless of flash-to-click delays, contrasting before-flip and after-flip click LFPs.

Data from two electrodes were illustrated as examples (**Fig. 2A**). In the example from Subject 1, click LFP waveform between 40- and 80-msec post-stimulus responses latency was negatively-shifted in the after-flip condition (the trace in red) when compared to the before-flip baseline (the trace in black). There was also a brief positive effect of visual modulation in LFP waveform of the same example at ~20-ms response. The difference marked by the yellow shading behind the waveforms were beyond a 95% statistical threshold ($\alpha = 0.05$, two-tailed), where a half of total fixation trials were resampled for 1000 times to estimate the probability of LFP difference greater or less than zero. We used the same method throughout this study for measuring the inter-trial consistency of any observed difference (see details in **Methods**). This index can be seen as a non-parametric form of z-score, and unlike z-score, it does not require data to conform to normal distribution. In the example from Subject 2, changes in LFP waveform were observed between 35- and 50-ms response latency. In this case, both the positive and the negative peak components on the LFP waveform were enlarged, that is, the positive component was positively-shifted while the negative component was vice versa (i.e., negatively-shifted).

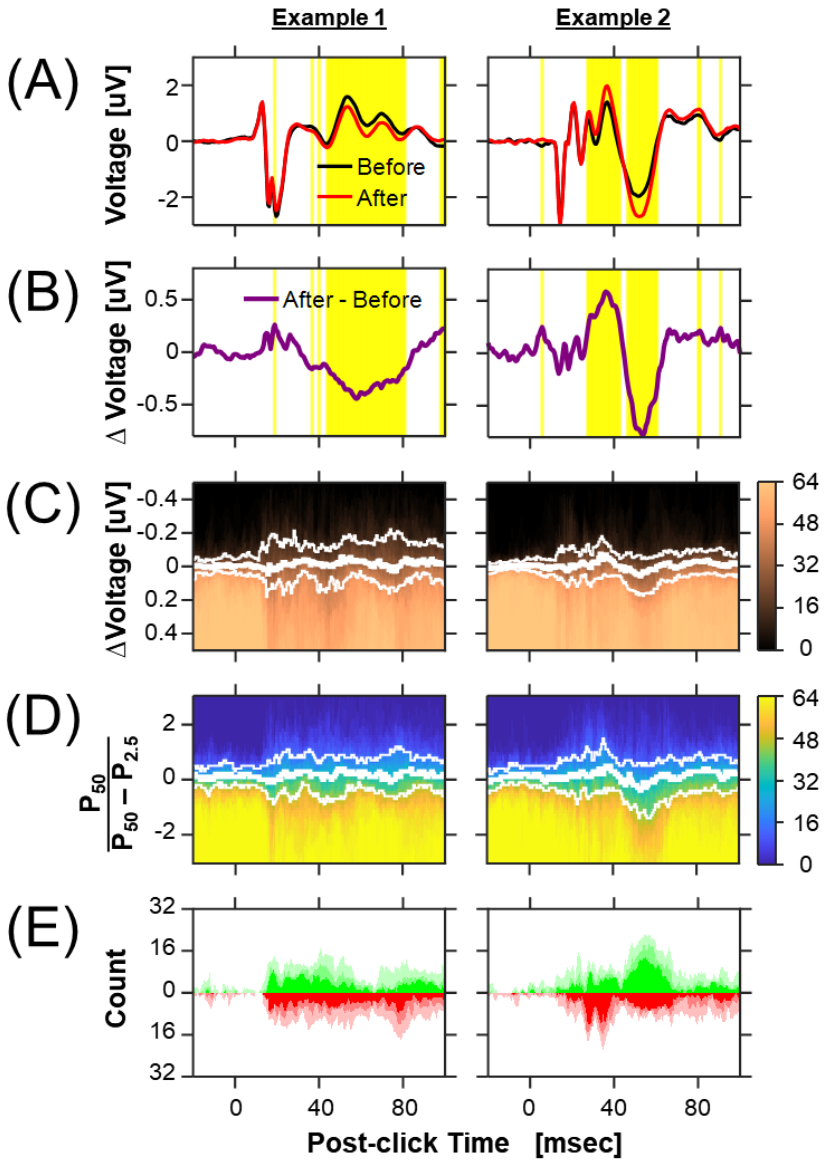


Figure 2. Effect of visual stimulus on click evoked local field potentials (LFPs) regardless of audiovisual disparity. (A) Examples of LFPs evoked by clicks before and after checkerboard flip in both subject. **(B)** LFP differences representing the effect of visual modulation derived from waveforms. Yellow shading, $p < .05$. **(C)** The cumulative distribution function (CDF) for LFP differences for each time point. Top, the count of electrodes with LFP difference less than the cut-off on the vertical axis is color-coded. Solid thick line, median values. Solid thin line, 25%- or 75%-quartile. **(D)** the CDF for an effect-size index based off inter-trial consistency for each time point. The count and quartile values were visualized using similar way. See Methods for the details on the calculation of this index. **(E)** the count of electrodes with positive (green) and negative (red) visual modulation beyond threshold of inter-trial consistency. Three different shadings from dark to bright represented for 0.1%, 1% and 5% thresholds (i.e., the alpha values).

Before the subsequent data analysis, we first subtracted the before-flip baseline from the after-flip LFP waveforms as a measurement of visual modulation (**Fig. 2B**). In the first example, the waveform of the LFP differences (ΔLFPs) did not appear to align in time with the click LFPs itself. The late-negative component of the ΔLFPs spanned through 3 negative and 2 positive click LFP components, with the largest negative ΔLFPs being closest to the second LFP peak. By contrast, in the second example, ΔLFPs aligned well in time with the click LFPs at the strongest response. Meanwhile, a positive change in LFPs was visualized at an early response latency (~7ms).

The population distribution of ΔLFPs was visualized by the cumulative electrode count with 25%, 50% and 75% quartiles highlighted (**Fig. 2C**). In Subject 1, the effect of visual modulation from the two 32-channel electrode matrices was spread across the majority of the response window (> 20 msec), including both positive and negative ΔLFPs. In Subject 2, negative ΔLFPs seemed to dominate the visual modulation of the recorded neurons in early responses (20-40 msec) while positive ΔLFPs, vice versa, dominated the late responses (40-70 msec). Next, we took the first 100 resampled ΔLFPs for deriving the median (i.e., P_{50}) and the 2.5 or 97.5 percentile ($P_{2.5}$ or $P_{97.5}$). The ratio between P_{50} and the margin of error, estimated by the distance between P_{50} and $P_{2.5}$ for positive P_{50}, or between P_{50} and $P_{97.5}$ for negative P_{50}, normalized ΔLFPs by its inter-trial variation, and thus can be used as an effect-size index for LFPs changes due to visual modulation (see details in **Methods**). The time course of the effect-size index of ΔLFPs was comparable to ΔLFPs and showed similar distribution of visual modulation recorded in both subjects (**Fig. 2D**).

Finally, in each subject we counted the number of electrodes where visual modulation was beyond thresholds of inter-trial consistency (**Fig. 2E**). The electrode counts overall had similar time course as the effect-size index and ΔLFPs. This analysis also showed that about 25% to 50% of electrodes demonstrated consistent visual modulation of click LFPs at certain response latencies, even though there was no temporal correspondence between the auditory and the visual stimuli.

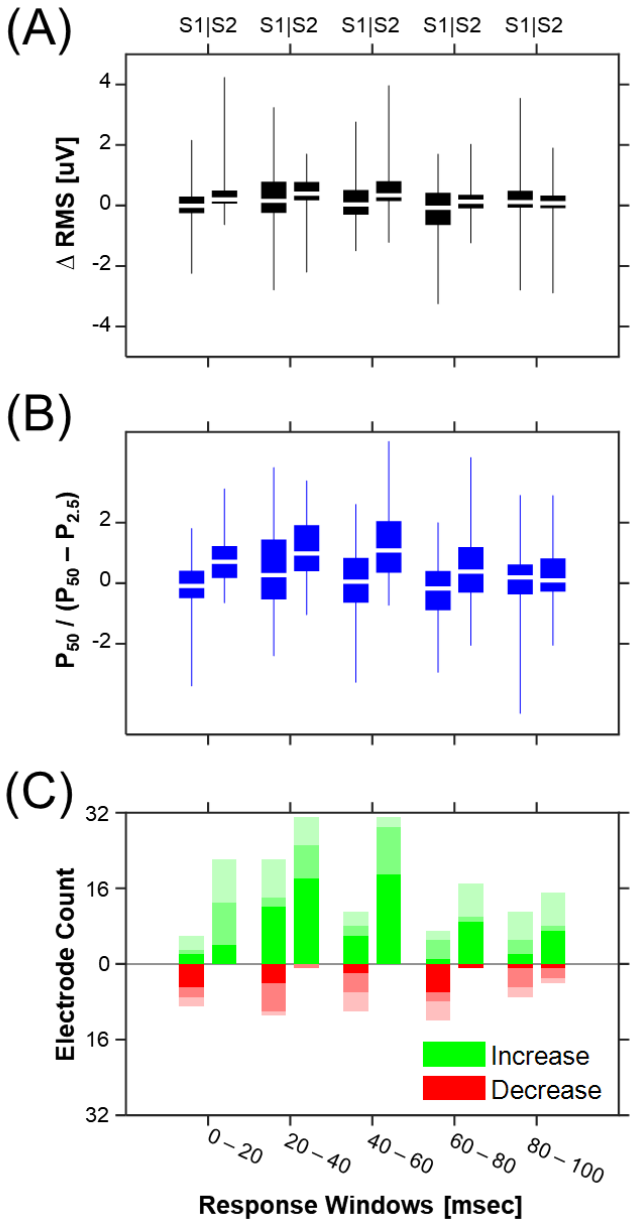


Figure 3. Effect of visual stimulus on signal power of LFP by response windows. (A) difference in root-mean-square (ΔRMS) values between LFPs evoked by clicks before and after checkerboard flip, which were calculated from each of the five response windows in both subjects. White line, median. Black box, inter-quartile range. Whisker, full range. S1, Subject 1. S2, Subject 2. **(B)** the effect-size index calculated for ΔRMS in the same way as for LFP differences. **(C)** the count of electrodes with ΔRMS beyond inter-trial consistency thresholds. Three different shadings from dark to bright represented for 0.1%, 1% and 5% thresholds (i.e., the alpha values).

To reduce the number of variables, we calculated root-mean-square (RMS) of LFPs in five 20-ms response windows as a measure of signal strength, which is equivalent to LFP power. Compared to an area-under-curve or rectified-voltage quantification (Ghazanfar et al., 2005), the calculation of RMS penalizes the sample points on the averaged LFP waveforms for small absolute magnitude, and therefore allows the synchronized neuronal activities time-locked to click onset to have a larger contribution to the RMS measurement. In both subjects, the RMS derived from after-flip LFPs demonstrated both increase and decrease from the before-flip baseline, with a large variation in their difference (ΔRMS) across electrodes (**Fig. 3A**).

Similar normalization procedure was performed on window-wise ΔRMS instead of sample-wise ΔLFPs. The effect-size index of ΔRMS revealed the effect of visual modulation with less inter-electrode variance (**Fig. 3B**). The upper quartile showed clearly a more prevalent RMS enhancement in the electrode population for the response windows between 20- and 40-msec in both subjects. In Subject 2, a trend of RMS enhancement was also seen for the response windows between 0- and 20-msec and between 40- and 60-msec.

Again, in each subject we counted the number of electrodes for ΔRMS beyond consistency thresholds (**Fig. 3C**). There were a number of electrodes in Subject 1 showing both significant enhancement and attenuation in RMS. Enhancement of RMS was more frequent for the middle response windows than the first and the last response windows. On the other hand, attenuation of RMS appeared more evenly-distributed across five response windows, and overall much less frequent than the enhancement. Statistical tests showed that the electrode counts were significantly different from zeros in all response windows for both enhancement and attenuation of RMS (**Supplementary Table 1**). In Subject 2, about 25%~50% of electrodes showed enhancement of

RMS by visual modulation, but only a couple of electrodes showed attenuation of RMS. Only the electrode counts for the enhancement were significantly different from zeros.

The effect of visual modulation with varying flash-to-click delays

To examine the effect of audiovisual disparity on the visual modulation in click LFPs, we first followed a classical approach where audiovisual disparities were sampled as a categorical variable. To do so, we grouped the after-flip clicks into 18 bins by their flash-to-click delays with each bin including the same number of clicks (**Supplementary Table 2**). Click LFPs were obtained for each group and compared to the before-flip baseline.

We also took three routes of characterizing the delay effect. First, we reported the occurrence of significant visual modulation for various delay and response window. This route of analysis demonstrated what delay conditions produce significant and reliable visual modulation of click LFPs. Second, we reported the modes of the optimal delay for the largest visual modulation in individual electrode (where mode refers to data with the most frequent occurrence), which demonstrated the most effective delay for relatively stronger visual modulation. Third, we counted the number of delay groups showing significant visual modulation, which evaluates the effective width of integration window.

In the example shown here (**Fig. 4A**), we stacked the waveforms of click LFPs from individual delay groups and observed highly consistent waveforms across delay groups. The effect of visual modulation, however, was not. For delays between 113- and 1269-msec, there was a larger negative component in the waveform of LFPs evoked by clicks at response latency of about 50-msec. The magnitude of this visual modulation varied with flash-to-click delay, with the largest

visual modulation for 727-msec delay group. The amount of ΔLFPs was almost negligible for some delay groups, such as 38-msec, 274-msec, 525-msec, and 838-msec delay groups.

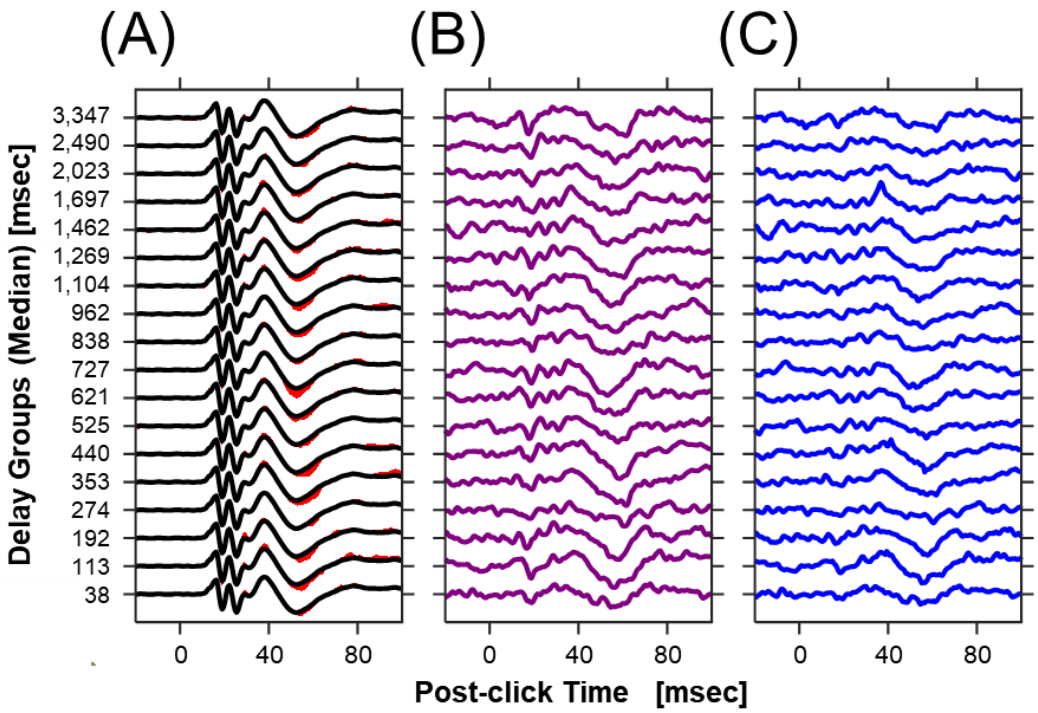


Figure 4. An example of visual modulation in click local field potentials (LFPs) by flash-to-click delay groups. (A) click evoked LFPs stacked by the median of flash-to-click delays from each group. The interval between two ticks on the vertical axis is 5 uV with the ticks aligned to zero of each LFPs. Black, baseline LFPs averaged from before-flip clicks. Red, gaps between baseline LFPs and LFPs of each delay group. **(B)** LFP differences stacked by the delays. The interval of the vertical axis is 2 uV. **(C)** the effect-size index for each time point stacked by the delays. The interval of the vertical axis is 5 (arbitrary unit).

In addition, the stacked waveforms of ΔLFPs revealed a separate component of visual modulation at an earlier latency around 20-msec (**Fig. 4B**). We again normalized ΔLFPs by its variation by calculating the effect-size index of ΔLFPs for each delay group (**Fig. 4C)**. The effect size of visual modulation after the normalization became less apparent than ΔLFPs themselves, but we were able to identify another positive visual modulation at ~35-ms response latency for the 1697-msec delay group, likely due to the low inter-trial variance in the LFPs of that group.

Therefore, this example is also a good demonstration that the effect-size index serves as a particularly useful measurement for comparison across different delay groups, as the fluctuation of ΔLFP variance across different delay groups is compensated.

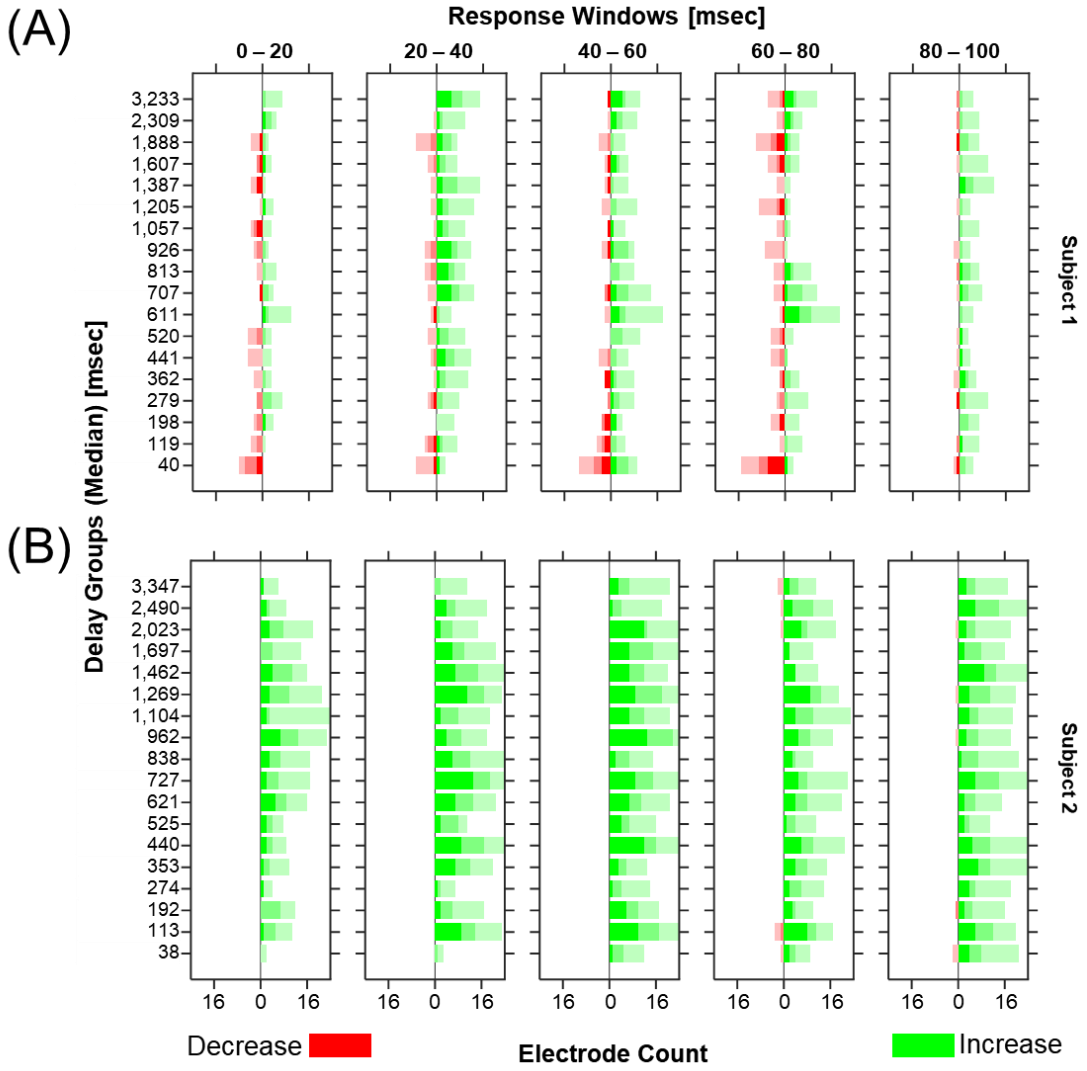


Figure 5. Effect of visual modulation in signal power of click LFPs by flash-to-click delay groups. The count of electrodes with ΔRMS beyond inter-trial consistency thresholds as a function of flash-to-click delay groups for each of five response windows in Subject 1 **(A)** and Subject 2 **(B)**. Three different shadings from dark to bright represented for 0.1%, 1% and 5% thresholds (i.e., the alpha values).

For the measurement of ΔRMS, we first visualized its population statistics for each delay group in our preliminary analysis (**Supplementary Fig. 1-2**). The medians of the effect-size index of ΔRMS, ranging from -0.28 to 0.56 in Subject 1 and from 0.33 to 1.03 in Subject 2, did not

demonstrate any contrasting difference either between delay groups or response windows. However, the count of electrodes beyond consistency thresholds (**Fig. 5**) clearly indicated a higher occurrence of RMS enhancement in certain delay groups. In Subject 1 (**Fig. 5A**), we observed more RMS enhancement for medium to long delays than short delays while RMS attenuation seemed to favor the shortest delay groups. The RMSs from Subject 2 (**Fig. 5B**) were enhanced in an overall larger proportion of electrodes than in Subject 1, and again, the RMS enhancement occurred more frequently at medium to long delay. Although the RMS attenuation was rare in Subject 2, they also occurred for the shortest delay groups.

To assess whether delay grouping helped with a higher chance of observing visual modulation in click LFPs, we ranked the delay groups by electrode counts in each subject, for increase or decrease in ΔRMS from five response windows separately. We found that the top five delay groups of each response window had comparable number of electrodes demonstrating visual modulations to those before delay grouping (**Supplementary Fig. 3**). When using the 5% consistency threshold, both subjects demonstrated more electrodes in RMS enhancement in the optimal delay groups when compared to the electrode counts without considering delays (**Supplementary Table 3)**. More electrodes in RMS attenuation with optimal delay groups were found in the response windows between 40- and 60-msec and between 60- and 80-msec for Subject 1, and between 60- and 80-msec for Subject 2. When using the 1% and the 0.1% consistency thresholds, slightly less electrodes were found with optimal delay groups. This is likely due to that, with a lower number of clicks in individual delay groups compared to the grand average, the effect of visual modulation was more likely to fail the statistical tests. Our findings suggested that it is possible to optimize the search of audiovisual interaction by sampling a broad range of audiovisual disparity for the identification of A1 neurons receiving visual modulation.

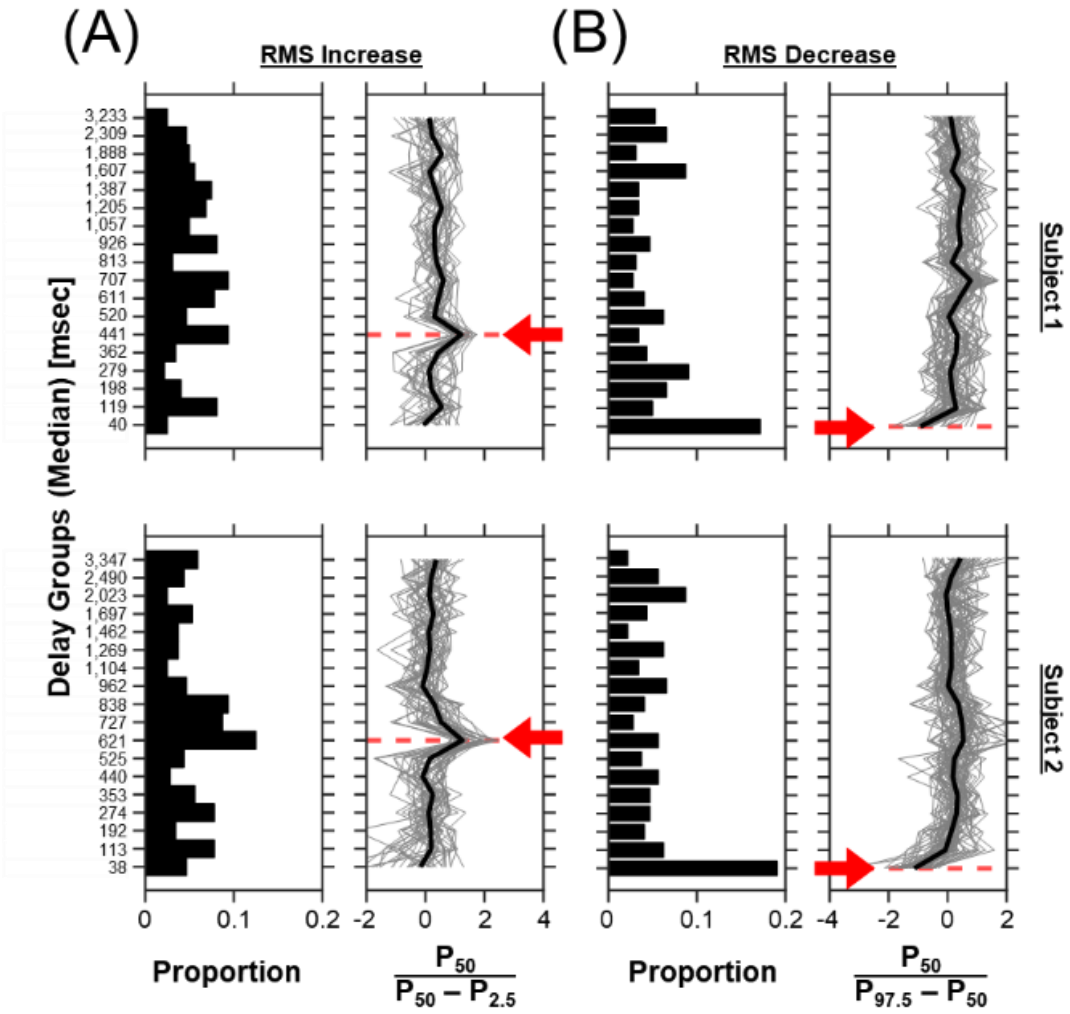


Figure 6. Distribution of optimal delays for visual modulation. The distribution of optimal delays with electrodes and response windows pooled together in Subject 1 (top) and Subject 2 (bottom). For given electrode and response window, the optimal delays were determined for the largest effect sizes of RMS enhancement **(A)** an RMS attenuation **(B)**. The delay functions that peaked at the optimal delay were demonstrated for each condition. Grey line, overlaid delay functions of effect-size index. Black, median of the overlaid delay functions.

We pooled all five response windows together and looked at the peak properties of individual delay functions of ΔRMS. We first plot the distributions of optimal delays in each subject for both RMS enhancement (**Fig. 6A, left**) and attenuation (**Fig. 6B, left**), which were defined as the only maximum or minimum out of 18 delay groups. In Subject 1, we found the mode of optimal delays for RMS enhancement to be the 441-msec group, with 707- and 119-msec

groups slightly behind. In Subject 2, the top three were 621-, 838- and 727-ms groups. In both subjects, the mode of optimal delay for RMS attenuation were found at the shortest delay group. We overlaid those delay functions of the optimal delay groups identified (**Fig. 6AB, right**). The medians of the overlaid delay functions were significantly different from the rest of the delay functions ($p < .001$) at the optimal delays. The distribution of optimal delays was also examined for each response window (**Supplementary Fig. 4-5**). Regardless of response windows, the distributions of RMS enhancement spread across various delay groups for different response windows while the distributions of RMS attenuation decrease were limited to the shortest delay groups. Altogether, our findings suggested that the response power of click LFPs is more likely to be attenuated by near-simultaneous visual stimulus and to be enhanced by a preceding visual stimulus for a wide-spread variation of temporal disparity.

It was also noticed that many delay functions appeared to show multiple peaks with several different delay groups (**Fig. 7A**). Therefore, we counted the number of delay groups with significant audiovisual interaction for each delay function and classified them into four classes including one-, two-, three-peak delay functions, and those with more than three. The multi-peak features appeared stronger for RMS enhancement (**Fig. 7B**), as indicated by more right-shifted frequency distributions in both subjects, than for RMS attenuation (**Fig. 7C**). Consistent with a previous finding on audio-tactile interaction in macaque A1, our results supported the hypothesis that neuronal multisensory interaction in A1 may involve multiple integration windows (Lakatos et al., 2007) and also suggested there may exist different mechanism for enhancement and attenuation of LFPs. Single integration window as previously discovered in midbrain (Meredith et al., 1987) and rat cortical (Schormans et al., 2017) audiovisual neurons may not apply to cat primary auditory cortex.

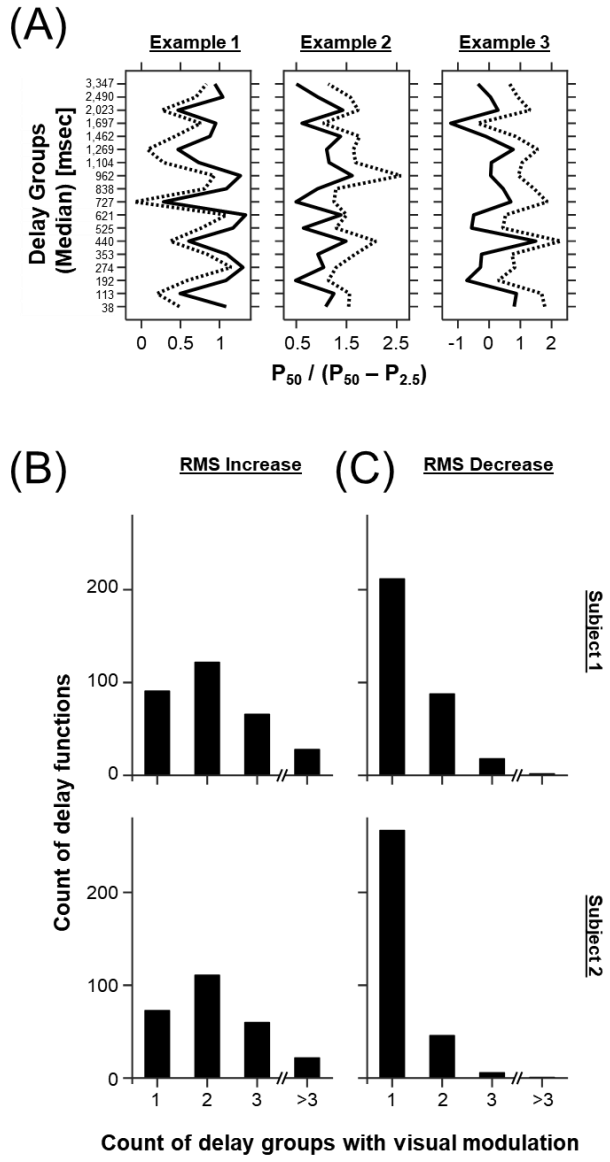


Figure 7. Occurrence of multiple effective delays within the same delay function. (A) Three delay functions as examples of multiple delay groups with equivalently effective visual modulation in ΔRMS. Solid, effect-size index of ΔRMS using before-flip clicks as baseline. Dotted, after-flip clicks as baseline. **(B and C)** The count of delay functions with 1, 2, 3, and >3 delay groups showing significant visual modulations were shown for Subject 1 (top) and Subject 2 (bottom).

Visual modulation of click LFPs depends on both response latency and audiovisual disparity

Although it is a common practise to characterize the effect of temporal disparity with a various number of selected delays on audiovisual interaction (Meredith, 1987; Kayser, 2008;

Schorman, 2018) or audio-tactile interaction (Lakatos, 2007) as we did in the previous section, the validity of this traditional approach can be vulnerably compromised by potential artifact and aliasing effects due to a low resolution of sampling in varying delays. To get around from this problem, we adopted a novel kernel-based approach in charactering the effect of temporal disparity. Briefly, we centered a Gaussian kernel function at varying temporal disparity with a 5-msec step and use the value of kernel function at the flash-to-click delay of each click as the weight coefficient of its response to derive the averaged LFP waveforms for each temporal disparity. In this way, each averaged LFP waveform receives more contribution from clicks with flash-to-click delay closer to the centered temporal disparity than those further away (see details in **Methods**).

We applied a similar statistical analysis to the LFPs derived using this approach as we did to the two examples in Figure 2. Visual modulation was quantified by ΔLFP and examined with the same resampling procedure for inter-trial consistency, with the evaluated temporal disparity spanning from 0 to 1500 msec. We visualized the count of electrodes with significant inter-trial consistency ($\alpha = .05$) as the shading of red (negative) or green (positive) in two-dimensional maps of audiovisual disparity (ordinates) and response latency (abscissa). In Subject 1, late response (40 – 80 ms) of click LFPs with small evaluated audiovisual disparity (< 200 ms) demonstrated visual modulation in the form of both negative (**Fig. 8A**) and positive ΔLFP (**Fig. 8B**), as well as a brief window of LFP at ~ 20 msec latency spanning a wide range of audiovisual disparity (300-1500 msec). In Subject 2, visual modulation of the highest electrode count resided in click LFPs at latency from 20- to 40-msec for negative ΔLFP or from 50- to 70-msec for positive ΔLFP. With a more conservative threshold of inter-trial consistency ($\alpha = .001$) (**Supplementary Fig. 6**), Subject 1 demonstrated more specificity for audiovisual disparity in terms of both positive and negative ΔLFP than Subject 2. A similar pattern was revealed by the population distributions of ΔLFP

across electrode and the effect-size index (**Supplementary Fig. 7**) as using electrode counts. Thereby, with the approach taken, we confirmed the observations of visual modulation in click LFPs for both subjects and displayed how this approach can be utilized to characterize the selectivity of visual modulation to response latency audiovisual temporal disparity.

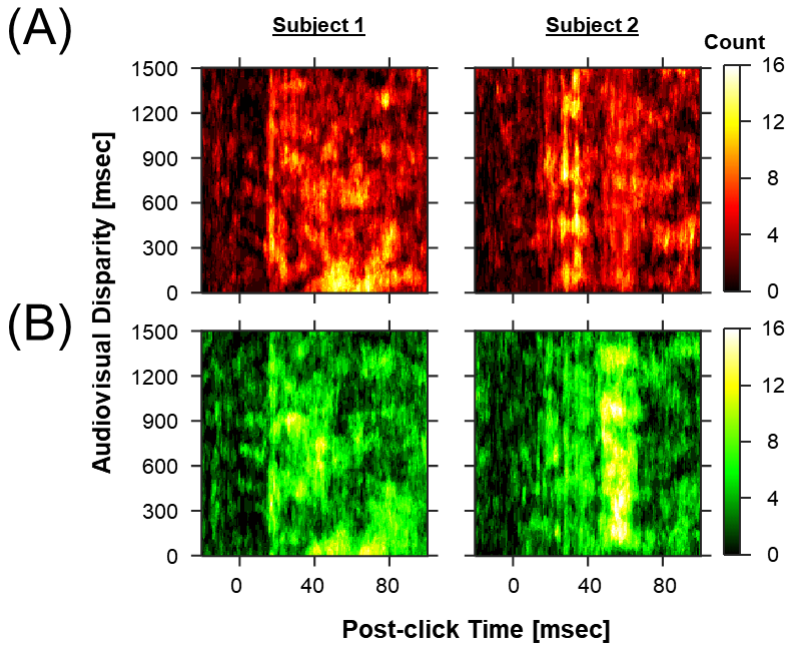


Figure 8. Estimated visual modulation as a function of audiovisual disparity. The count of electrodes with visual modulation beyond inter-trial consistency threshold (5%) for each time point in both subjects. LFPs estimated for different audiovisual disparities on the vertical axis were compared to the same before-flip baseline. Negative changes **(A)** and positive changes **(B)** were separated.

Similar as in one-dimensional delay functions constructed with RMS values, one or multiple clusters of response latency and temporal disparity were effective for significant visual modulation (**Supplementary Fig. 8**). After removing minor clusters of a size less than 500 ms-by-ms, we again classified total electrodes into four categories by the criteria of demonstrating one, two, three, or more than three major parameter clusters for visual modulation. In Subject 1, the largest class of electrodes were mono-cluster for positive ΔLFPs, which roughly tripled those of

dual-cluster or trio-cluster. For negative ΔLFPs, there were about twice more mono- and trio-cluster electrodes than duo-cluster electrodes. Electrodes showing more than three parameter clusters for visual modulation in both cases were rare. In Subject 2, although there were still less trio-cluster electrodes than mono- or duo-cluster electrodes, the difference was much less contrasting than in Subject 1. Again, the results of this analysis were consistent with the previous findings of multiple temporal integration windows for multisensory interaction revealed at neuronal level (Lakatos et al., 2007).

Finally, we characterized the optimal combinations of response latency and temporal audiovisual disparity at population level as the probability density function of local extremes (maxima and minima) in ΔLFPs using kernel density estimation (**Supplementary Fig. 9**). The estimated modes of response latencies for local extremes lied somewhere between 20- and 80-msec, depending on the subjects and the directions. In Subject 1, the probability density of ΔLFPs maxima peaked at ~700-msec temporal disparity, while the probability density of ΔLFPs minima peaked at ~100-msec delay. In Subject 2, the probability density of ΔLFPs minima peaked at ~400-msec delay, while the probability density of ΔLFPs maxima remained on a plateau throughout almost the entire range of evaluated audiovisual disparity.

Periodic patterns in the temporal course of visual modulation of click LFPs

It has been proposed that multiple temporal integration windows are related to the oscillations of excitability coupled to frequencies in different EEG bands (Lakatos et al., 2007). The previous study has shed some light on this issue by showing that the optimal stimulus delay within temporal integration windows coincided with the interval of one cycle of frequency-specific brain oscillation. Led by our findings of multiple integration windows, we further scrutinized the

presence of periodic patterns in the dynamics of visual modulation as a function of audiovisual disparity.

We revisited this question using LFPs data and the current novel stimulus paradigm with an approach that can be seen as frequency decomposition. Similar data analysis was adopted in a few human studies (Diederich, Schomburg, & Colonius, 2012; Romei, Gross, & Thut, 2012) to characterize the dynamics of excitability in visual cortex modulated by a preceding auditory stimulus in the frequency domain. To avoid any bias from prior knowledge, we selected discrete frequencies from 1 Hz to 80 Hz at 1-Hz step to construct a "spectrum" of arbitrary magnitudes based on a phase-lock measurement. A similar weight-average process was carried out, with the previous Gaussian function replaced by a modified Cosine function (see **Methods**). We obtained 12 phase-shifted versions of visual modulation quantfied as ΔRMS and performed the statistical analyses on their vector sum (i.e., resultant) for each frequency.

It was unexpected to find in both subjects that the most prominent periodic visual modulation effect was at 60 Hz for RMS from the first three response windows, i.e., between 0- and 20-msec, 20- and 40-msec, and 40- and 60-msec (**Fig. 9**), given that the extracellular recording was already notched-filtered at 60 Hz in the stage of data preprocessing. We considered this 60-Hz effect to be likely resulted from an unintended frequency entrainment associated to the frame rate of visual stimuli presented on the LCD screen, as it was shown previously that human and macaque visual cortex could be entrained to the refresh rate of CRT screen (Williams, Mechler, Gordon, Shapley, & Hawken, 2004). Frequency entrainment in the visual cortex has been reported to modulate auditory excitability in both human (Gomez-Ramirez et al., 2011) and animal (Lakatos, Karmos, Mehta, Ulbert, & Schroeder, 2008) subjects. The magnitudes of this 60-Hz effect was not uniformly distributed over the five response windows, suggesting that the effect of

frequency entrainment initiated from the visual inputs interacted with auditory processing of clicks, rather than the spontaneous neural activities in A1.

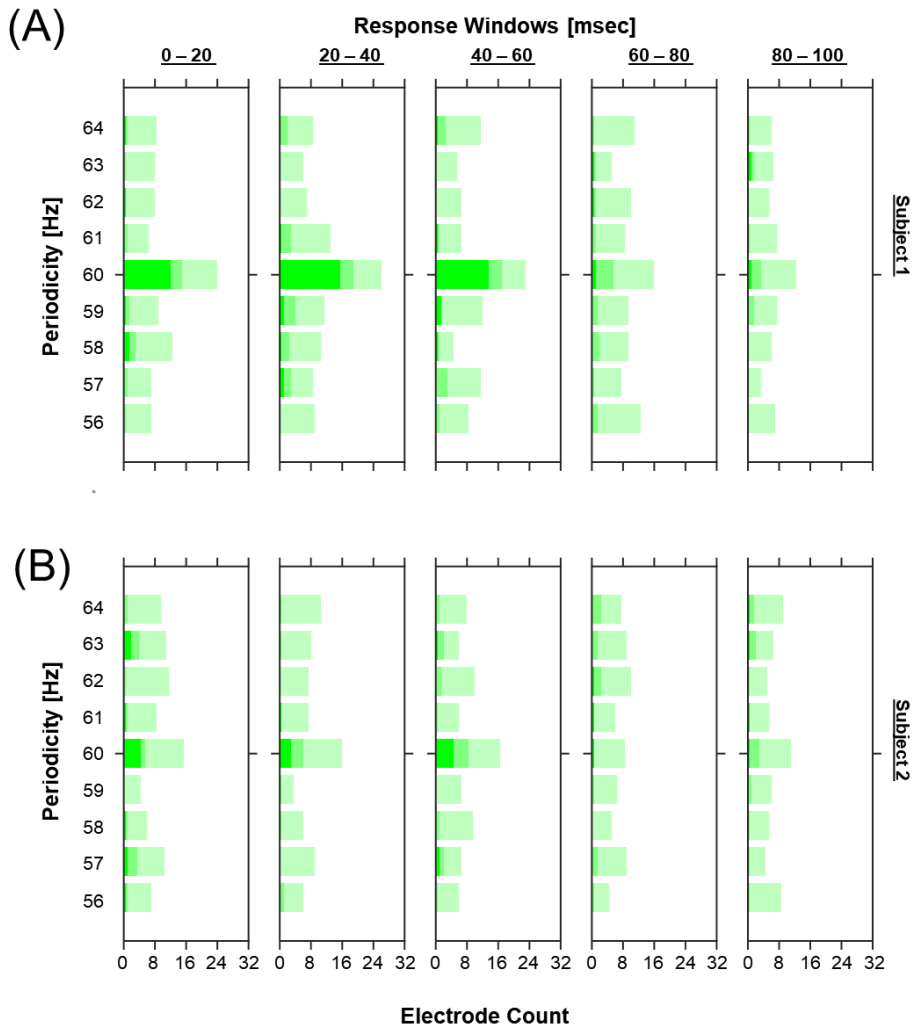


Figure 9. Frequency coupling of visual modulation in signal power of click LFPs. A 60-Hz effect shown by the count of electrodes with phase-locking amplitude beyond inter-trial consistency thresholds in Subject 1 **(A)** and Subject 2 **(B)**.

We did not observe phase-locking effect as strong as at 60-Hz for any other frequencies, which was not surprising considering the visual stimulus was not rhythmic. We did observe slightly more electrodes showing significant phase-locking visual modulation at the lowest

frequencies (1-4 Hz) than other frequencies for Subject 1, specific to the middle response windows (**Supplementary Fig. 10-11**). To reduce the variance of electrode count, we combined frequencies into 6 bands using conventional cut-off frequencies for EEG signal, with the delta band from 1 Hz to 4 Hz, the theta band from 5 Hz to 8 Hz, the alpha band from 9 Hz to 12 Hz, the beta band from 13 Hz to 30 Hz, low gamma band from 31 Hz to 59 Hz, and high gamma band from 61 Hz to 80 Hz **(Fig. 10)**. After the combination, the effect of frequency coupling in Subject 1 **(Fig. 10A)** became particularly prominent in the delta band (1-4 Hz). In Subject 2 **(Fig. 10B)**, slightly more electrodes demonstrated significant phase-locking visual modulation in the delta band than the theta and the alpha bands for the response window from 0- to 20-msec and from 20- to 40-msec. Our results suggested that there exists periodic pattern of varying frequencies in the dynamics of visual modulation in click LFPs of A1 neurons. In particular, the periodicity of delta-band frequency, potentially derived from the delta wave EEG across multiple sensory cortices, may have a unique contribution to the temporal course of cortical cross-modal interaction.

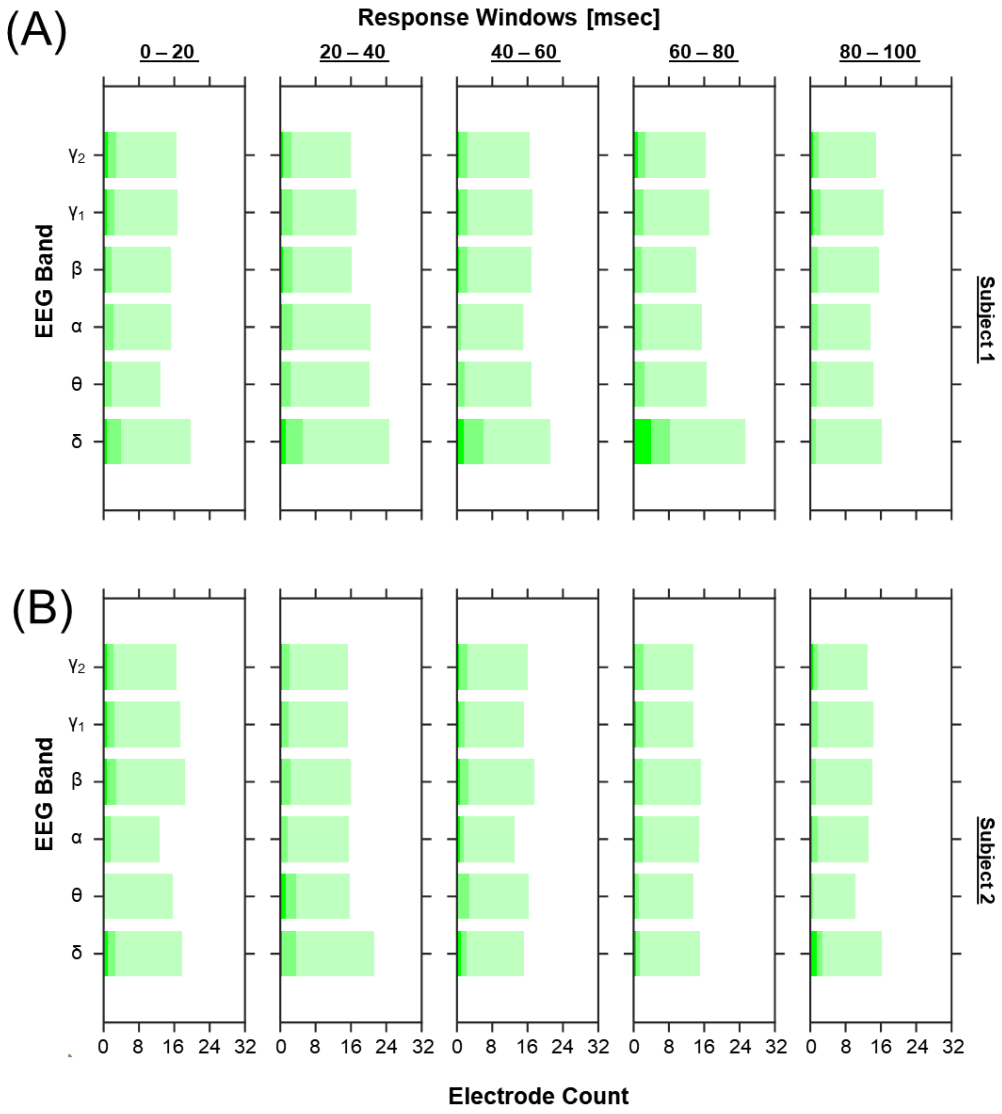


Figure 10. Frequency coupling of visual modulation for different EEG bands. The count of electrodes with phase-locking amplitude beyond inter-trial consistency thresholds averaged for frequencies from different EEG bands in Subject 1 **(A)** and Subject 2 **(B)**. Three different shadings from dark to bright represented for 0.1%, 1% and 5% thresholds (i.e., the alpha values).

3.5. Discussion

Characterizing the effect of stimulus timing in visual modulation with delay functions

Visual modulation in auditory neurons has been studied in the primary auditory cortex in various animal models, including cats (Kral et al., 2003; Rebillard et al., 1977; Stewart & Starr, 1970), non-human primates (Ghazanfar et al., 2005; Kayser et al., 2008; Lakatos et al., 2007), ferrets (Bizley et al., 2007; Meredith & Allman, 2012, 2015), mouse (Hunt, 2006), but most of them used a fixed temporal disparity in audiovisual stimuli, except for Kayser (2008) and Lakatos (2007). Even for multisensory studies investigating beyond the primary auditory cortex, only a few have investigated the temporal disparity as a factor (Meredith et al., 1987; Schormans et al., 2017).

Meanwhile, behavior and brain imaging studies in human subjects have widely adopted audiovisual delay as a factor (Miller & D'esposito, 2005; Ross, Molholm, Butler, Del Bene, & Foxe, 2022; Stevenson, Altieri, Kim, Pisoni, & James, 2010; van Atteveldt, Formisano, Blomert, & Goebel, 2007), although most of them selected the test range of audiovisual temporal disparity based on its behavioral relevance, i.e., perceived simultaneity. However, it has been recently shown that the cross-modal enhancement of perceptual performance was independent of the judgement of audiovisual simultaneity (Harrar, Harris, & Spence, 2017). As was shown in a human behavior study (Fiebelkorn et al., 2011) and our results, cross-modal interaction was not necessarily absent for a large audiovisual temporal disparity that is already outside of the possible range of perceived simultaneity. Experimental data with stimulus conditions of such long temporal disparity are still understandably scarce in existing literature, due to the prolonged recording duration in order to contain a huge number of trials needed for various audiovisual delays when using a classical two-impulse paradigm. Therefore, one aim of the current study was to explore a novel stimulus paradigm that allows for more efficient data collection in characterizing the effect of temporal disparity. By obtaining click LFPs from selected flip-to-click delays, we characterized

the relationship between visual modulation and audiovisual temporal disparity. The "delay function" of visual modulation quantified the effect of checkerboard flip on how neuronal excitability in cat A1 dynamically evolved over time (i.e., the temporal course).

In this study, the temporal course of visual modulation at individual electrode showed that optimal delay for enhanced signal strength (RMS) of click LFPs was not necessarily within the range for perceived audiovisual simultaneity but rather varied substantially from electrode to electrode. The optimal delay for RMS attenuation, however, was found at the smallest delay group (with a median of flip-to-click delays about 40-ms) more frequently than other delay groups. The exact neural circuits that mediate the delay-tuning features for cross-modal interaction is still unclear, although studies in auditory temporal processing have frequently attributed a temporal pattern of neural activities to the synergistic outcome of a "push-and-pull" process from excitatory and inhibitory inputs (Brosch & Schreiner, 1997; Joris & Yin, 2007; Pienkowski, Shaw, & Eggermont, 2009). One such example arises from midbrain binaural neurons, where response tuning to the interaural time difference (i.e., ITD) at a sub-millisecond time scale was shaped by both depolarizing sodium currents as excitatory inputs and glycine-mediated inhibitory inputs (Brand, Behrend, Marquardt, McAlpine, & Grothe, 2002; Zhou, Carney, & Colburn, 2005).

It is also worth mentioning that recent behavioral studies have shed some light on the plasticity in cross-modal temporal processing (Fujisaki, Shimojo, Kashino, & Nishida, 2004; Keetels & Vroomen, 2007; Navarra et al., 2005; Powers et al., 2009; Vroomen, Keetels, De Gelder, & Bertelson, 2004). After subjects were exposed to auditory- or visual-leading audiovisual stimuli, the delay of perceived simultaneity shifted in the corresponding direction. In gerbils, it was shown that exposure to the asynchronized audiovisual stimuli changed the directional influences between the auditory cortex and the visual cortex (Fillbrandt & Ohl, 2012). Since the largest audiovisual

lag for effective exposure was found up to 600-ms (Fujisaki et al., 2004), it is possible that the primary auditory cortex is wired in the first place for processing long audiovisual temporal disparity, and, perhaps, mediating the underlying neural plasticity.

A crucial discrepancy in methodology between the current study and the previous reports regarding the delay functions of multisensory interaction lies on the temporal context (i.e., background) of auditory stimuli. In our stimulus paradigm, the subjects were exerted no bias towards perceptual binding in any pair of click and checkerboard flip. Such bias was inevitable in a classical two-pulse stimulus paradigm, which is in the context of silent background. It has been shown in anesthetized ferrets that sensitivity of A1 neurons to the stimulus deviance was increased when the deviance-contained auditory stream bound with a co-present visual stream temporally (Atilgan et al., 2018), which suggests a strong bottom-up effect of binding driven by audiovisual stimuli. It cannot be excluded that the preference of 40- to 80-ms audiovisual temporal disparity previously reported in macaque and rat auditory cortex (Kayser et al., 2008; Schormans et al., 2017) is associated to this binding effect.

Visual modulation of auditory activation without temporal correspondence

Audiovisual temporal correspondence is known to serve as a strong promotor in the processing of speech (Stevenson et al., 2010; van Atteveldt et al., 2007). One study found that a co-present uncorrelated train of flash pulses suppressed BOLD signals in response to a train of clicks in voxels residing in superior temporal sulcus and the auditory cortex, in comparison to auditory-only baseline (Noesselt et al., 2007). Using the same stimulus paradigm, a behavior study showed that the performance of target detection was undermined by unsynchronized audiovisual stimulus (Marchant, Ruff, & Driver, 2012). In contrast, another behavioral study

showed that there was still multisensory benefit without correlation between auditory and visual stimuli when the task was to compare the overall rates of pulses (Raposo, Sheppard, Schrater, & Churchland, 2012). The stimuli used in the current study, which are made of clicks with stochastic flash-to-click delays, are very similar to the asynchronized audiovisual stimuli used in those studies except for a reduced rate in visual pulse train. The observations of visual modulation when contrasting click LFPs present after checkerboard flip with before-flip baseline are in support of the hypotheses that there may exist a separate route of audiovisual integration for the processing of only gross information (e.g., duration), where the exact temporal information is discarded (Raposo et al., 2012).

Frequency coupling in the temporal course of visual modulation

The classic shape of delay function of either neural or behavioral measurements, which appears like a single peak near zero audiovisual delay (or known as bell-shaped), has been treated as one of the fundamental properties in the temporal processing of audiovisual stimulus for 35 years even since it was first discovered (Meredith et al., 1987). It was, however, hard to reconcile with the observation of multiple temporal integration windows (Lakatos et al., 2007). According to the phase reset hypothesis the authors proposed, the cortical excitability indicated by click LFPs was modulated by EEG oscillations. Therefore, the optimal audiovisual temporal disparity corresponds to the interval of each oscillation cycle periods, whereas the initial phase of the oscillations is time-locked to checkerboard flips.

It has also been implied that, when visual stimuli are presented in a rhythmic manner, it is possible to enhance the cortical excitability featuring increased amplitude on the spectrum of EEG near the frequency of stimulus rhythm. The effect of rhythmic stimuli has already been revealed

in both human (Giraud & Poeppel, 2012) and macaque (O'Connell et al., 2015), known as frequency entrainment. Our data also demonstrated such frequency coupling in neuronal excitability in A1 time-locked to a single flip of checkerboard, in particular for delta band frequencies. This observation could be interpreted as an intrinsic property of neural circuit involving A1, although its exact function and mechanism is unknown. It is possible that A1 neurons that are set ready by a single checkerboard flip will have a frequency-specific preference for frequency entraining, should there be more upcoming checkerboard flips or clicks present to the subjects rhythmically.

Using behavioral measurements, there have been several studies in human subjects that investigated the periodic pattern of how a single-pulse stimulus in one modality affects the excitability of a following stimulus in another modality (Diederich et al., 2012; Diederich, Schomburg, & Van Vugt, 2014; Fiebelkorn et al., 2011; Romei et al., 2012). However, a unanimous agreement on the exact spectral pattern was not reached yet. The current study is among one of the very first research showing at neuronal level that a single-pulse, brief visual stimulus can modulate the response of following auditory stimulus with periodicity of varying frequencies. The stimulus paradigm we used makes it distinctive from the previous studies where the acoustic inputs itself is deposited with spectral information such as in narrowband noise and speech (Atilgan et al., 2018; Luo, Liu, & Poeppel, 2010). The auditory cortex is capable of extracting and tracking such spectral information in auditory stimulus. Although more data is needed to determine its variation across different A1 neurons, we proposed that the frequency coupling measured with our audiovisual stimuli may provide a new approach to characterize the stimulus-independent properties of neural circuits.

3.6. Methods

Animal preparation and experiment apparatus

All procedures were conducted in compliance with the National Research Council's Guide for the Care and Use of Laboratory Animals (8th edition; 2011) and the Canadian Council on Animal Care's Guide to the Care and Use of Experimental Animals (1993. Furthermore, the following procedures were also approved by Animal Care Committee (DOWB) for the Faculty of Medicine and Health Sciences at McGill University.

Once subjects reached a proficiency in performing a fixation task (while stimulus is present passively), they were chronically implanted with two recording electrode matrices in the left primary auditory cortex (A1). Approximately one-third of A1 area was covered by each electrode matrix, away from major blood vessels. One subject was also implanted with a single-needle metal electrode at the left medial genicular body for another research project.

During training and recording sessions, the subjects were seated on the top of a platform, wearing a loose veterinary restraining bag made of canvas fabric that was hanged on the vertical pillars on both sides. The subjects were not head-restrained. A 27-inch LCD screen (XL2720, BenQ) with 59.7-cm width and 23.5-cm height was placed in front of subjects at about 30-cm distance on the same horizontal plane of subjects' eyes. The estimated visual angle of the full-screen stimulus is about ±46-degree width and ±22-degree height. A speaker (MF1, TDT) with 5-cm diameter was placed underneath the screen on the midline. Between the subject and the screen, a plastic kitchen spoon was placed below the height of the speaker for dispensing canned cat food as reward at the end of success trials. The spoon was bottom-filled by a 60-ml syringe with catheter tip, which was pressed by a programmable linear electrical actuator. Above the subject's head was

an analogue passive head-stage amplifier (TDT) connected to the recording electrode matrices through an adaptor (Neuronexus or TDT), with its cable suspended and connected to the battery-powered digitizer (PZ5, TDT) placed nearby. The entire apparatus was contained in a 3-m-by-3-m double shielded acoustic chamber (IAC).

Visual and auditory stimuli

Each trial started with a visual cue to prompt the subject for the initiation of fixation. This visual cue was made of a 1-deg white dot centered on the screen flashing at 2 Hz. The background of the white dot was fully occupied by a 32-by-18 checkerboard made of 20-lux bright cells and 0.27-lux dark cells. Once the subject initiated the fixation at it, the flashing dot became steady and remained in the steady status until the subject's movement is detected by experimenter or a camera-based movement detection algorithm, which marks the end of a trial. When the fixation dot is in the steady status, a train of 100-us click pulses were played at 65 dB SPL. The timing of click pulses was determined by a Poisson random process with 20-ms deadtime between two consecutive clicks. On average clicks were played at a rate of 6.8 per second (Subject 1) and 6.9 per second (Subject 2). For the visual stimulus, the checkerboard at the background flipped to its invert version every time after the subject remained in the fixation status for ~2 seconds since the beginning of the fixation or the previous checkerboard flip. The exact interval between two flips was also randomized. On average the checkerboard flipped at a rate of 0.48 per second (Subject 1) and 0.43 per second (Subject 2).

Extracellular recording and local field potentials

Each of two microelectrode matrices (FMA, MicroProbes for Life Science) was made of 32 Platinum-Iridium electrodes (4-by-9, including 4 ground electrodes) wired perpendicularly to

a 1.8-mm-by-4.0-mm alumina ceramic platform. The depth of individual electrode is customized to 0.9-, 1.2-, or 1.5-mm. Signals of 64 channels in total were sampled at ~24.4 kHz (PZ5, TDT) and stored on a computer hard drive for offline data analysis.

Data were collected from each subject over 10 to 15 days, with one recording session of about 1-hour long for each day. The raw signal for each recording session was first pre-processed by passing a notch filter at 60 Hz and then a 10 Hz – 3000 Hz bandpass filter, and down-sampled at a ratio of 1:32. Signal epochs between two trials, or in trials with artifact (>500 uV) or trials too short (< 2.5 seconds) were discarded. The pre-processed signal had 64 channels in total for each of 32 electrodes from two matrices and went through a principal component analysis (PCA). Further data analysis was performed on the signal from individual channels with the contribution of the first 5 PCs subtracted.

To remove visually evoked potentials (VEPs), signal epochs from 100-ms before to 500-ms after checkerboard flip were averaged for each recording session. The VEP waveforms were baseline corrected and with its first and last 50-ms multiplied by a Hanning window, before it was subtracted from each individual trial. Click evoked local field potentials (LFPs) were derived from the average of signal epochs from 20-ms before to 100-ms after click onset, with each epoch baseline corrected.

On the other hand, flip-to-click delays were calculated for each individual click as the lag of its onset to the checkerboard flip preceding to it. Click present before checkerboard flip were used to estimate click LFPs as a baseline in auditory-only condition. In time-binned averaging, all click epochs were sorted in ascending order by flip-to-click delays and binned for about every 1000 epochs, which were averaged to derive the click LFPs for each bin or, namely, delay group.

In Gaussian-weight averaging, click LFPs for varying audiovisual disparity were sampled for delays from 0 to 1500-ms with a 5-ms step. For each sample, a Gaussian kernel function (σ = 50 ms) was centered at a given delay, and the click epochs with flip-to-click delays within a range of $\pm 3\ \sigma$ were included into the average. The weight of individual epoch was given by the output of the Gaussian kernel function for the delay with its flip-to-click delays as input. In this way, the averaged click LFPs are less contributed by click epochs with flip-to-click delays that are more distant from the peak of Gaussian kernel or the given delay, and vice versa. In cosine-weight average, the role of Gaussian kernel function was replaced by a cosine function with its negative values set to zero. For testing each frequency from 1 to 80 Hz at 1-Hz step, click epochs were averaged with weights given by the output of the modified cosine function. This was repeated for 12 times where in each time the initial phase of the cosine functions was shifted with 30 degrees.

RMS measurements and statistics

Signal strength was measured using root-mean-square (RMS) for five consecutive 20-ms response windows after click onset, which were from 0- to 20-ms, from 20- to 40-ms, from 40- to 60-ms, from 60- to 80-ms, and from 80- to 100-ms in response latency. Difference in click LFPs or RMSs in each response windows in compared to the before-flip baseline were measured as the effect of visual modulation. The complex mean or the resultant vector of the 12 phase-shifted versions of click LFPs determined the optimal phase of frequency coupling between visual stimulus and click LFPs (or auditory excitability).

To estimate the inter-trial consistency of visual modulation in click LFPs and RMS measurements, a 1000 subsampling of a half of the trials were randomly selected to derive additional samples used for estimating the distribution of inter-trial variations. An additional step

was taken for the RMS values from the Cosine-weight average, where the complex means derived from subsampling were projected to the angle of the optimal phase of frequency coupling previously calculated with all the trials. Samples of the ΔLFP, ΔRMS, or the optimal phase projection greater than zero were counted as a proportion out of 1000, and compared with three cut-off values (α), which were 0.001, 0.01, and 0.05 (two-tailed). The median and the 2.5^{th} or 97.5^{th} percentile were estimated from the first 100 subsampling, instead of the total 1000 subsampling to save calculation time. For positive visual modulation where the median was greater than zero, the ratio between the median and the distance between median and the 2.5^{th}-to-median distance was calculated as an effect-size index. For negative visual modulation, the 2.5^{th}-to-median distance was replaced by the median-to-97.5^{th} distance.

Fisher's exact test determines that an electrode count of 6 out of 64 in total yields a false-positive rate (i.e., P-value) of 0.0276. When Bonferroni correction is applied, an electrode count of 10 out of 64 yields a P-value of 0.0013, less than 0.0028 (0.05 divided by 18). Chi-square tests were performed to test whether the optimal delays followed a uniform distribution and whether the count of multi-peak delay functions followed a Poisson distribution. All data analysis was performed using Matlab and its toolboxes.

3.7. References

Atilgan, H., Town, S. M., Wood, K. C., Jones, G. P., Maddox, R. K., Lee, A. K., & Bizley, J. K. (2018). Integration of visual information in auditory cortex promotes auditory scene analysis through multisensory binding. *Neuron, 97*(3), 640-655. e644.

Bizley, J. K., Nodal, F. R., Bajo, V. M., Nelken, I., & King, A. J. (2007). Physiological and anatomical evidence for multisensory interactions in auditory cortex. *Cerebral Cortex, 17*(9), 2172-2189.

Brand, A., Behrend, O., Marquardt, T., McAlpine, D., & Grothe, B. (2002). Precise inhibition is essential for microsecond interaural time difference coding. *Nature, 417*(6888), 543-547.

Brosch, M., & Schreiner, C. E. (1997). Time course of forward masking tuning curves in cat primary auditory cortex. *Journal of Neurophysiology, 77*(2), 923-943.

Butler, B. E., Chabot, N., & Lomber, S. G. (2016). Quantifying and comparing the pattern of thalamic and cortical projections to the posterior auditory field in hearing and deaf cats. *Journal of Comparative Neurology, 524*(15), 3042-3063.

Calderone, D. J., Lakatos, P., Butler, P. D., & Castellanos, F. X. (2014). Entrainment of neural oscillations as a modifiable substrate of attention. *Trends in Cognitive Sciences, 18*(6), 300-309.

Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C., McGuire, P. K., . . . David, A. S. (1997). Activation of auditory cortex during silent lipreading. *Science, 276*(5312), 593-596.

Diederich, A., Schomburg, A., & Colonius, H. (2012). Saccadic reaction times to audiovisual stimuli show effects of oscillatory phase reset. *PloS One, 7*(10), e44910.

Diederich, A., Schomburg, A., & Van Vugt, M. (2014). Fronto-central theta oscillations are related to oscillations in saccadic response times (SRT): an EEG and behavioral data analysis. *PloS One, 9*(11), e112974.

Dixon, N. F., & Spitz, L. (1980). The detection of auditory visual desynchrony. *Perception, 9*(6), 719-721.

Fiebelkorn, I. C., Foxe, J. J., Butler, J. S., Mercier, M. R., Snyder, A. C., & Molholm, S. (2011). Ready, set, reset: stimulus-locked periodicity in behavioral performance demonstrates the consequences of cross-sensory phase reset. *Journal of Neuroscience, 31*(27), 9971-9981.

Fillbrandt, A., & Ohl, F. W. (2012). Effects of prolonged exposure to audiovisual stimuli with fixed stimulus onset asynchrony on interaction dynamics between primary auditory and primary visual cortex. *The Neural Bases of Multisensory Processes*.

Finney, E. M., Clementz, B. A., Hickok, G., & Dobkins, K. R. (2003). Visual stimuli activate auditory cortex in deaf subjects: evidence from MEG. *Neuroreport, 14*(11), 1425-1427.

Finney, E. M., Fine, I., & Dobkins, K. R. (2001). Visual stimuli activate auditory cortex in the deaf. *Nature Neuroscience, 4*(12), 1171-1173.

Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. y. (2004). Recalibration of audiovisual simultaneity. *Nature Neuroscience, 7*(7), 773-778.

Ghazanfar, A. A., Maier, J. X., Hoffman, K. L., & Logothetis, N. K. (2005). Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *Journal of Neuroscience, 25*(20), 5004-5012.

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience, 15*(4), 511-517.

Gomez-Ramirez, M., Kelly, S. P., Molholm, S., Sehatpour, P., Schwartz, T. H., & Foxe, J. J. (2011). Oscillatory sensory selection mechanisms during intersensory attention to rhythmic auditory and visual inputs: a human electrocorticographic investigation. *Journal of Neuroscience, 31*(50), 18556-18567.

Harrar, V., Harris, L. R., & Spence, C. (2017). Multisensory integration is independent of perceived simultaneity. *Experimental Brain Research, 235*, 763-775.

Hunt, D., Yamoah, E., & Krubitzer, L. (2006). Multisensory plasticity in congenitally deaf mice: how are cortical areas functionally specified? *Neuroscience, 139*(4), 1507-1524.

Joris, P., & Yin, T. C. (2007). A matter of time: internal delays in binaural processing. *Trends in Neurosciences, 30*(2), 70-78.

Kayser, C., Petkov, C. I., & Logothetis, N. K. (2008). Visual modulation of neurons in auditory cortex. *Cerebral Cortex, 18*(7), 1560-1574.

Kayser, C., Petkov, C. I., & Logothetis, N. K. (2009). Multisensory interactions in primate auditory cortex: fMRI and electrophysiology. *Hearing Research, 258*(1-2), 80-88.

Keetels, M., & Vroomen, J. (2007). No effect of auditory–visual spatial disparity on temporal recalibration. *Experimental Brain Research, 182*, 559-565.

Kobayasi, K. I., & Riquimaroux, H. (2013). Audiovisual integration in the primary auditory cortex of an awake rodent. *Neuroscience Letters, 534*, 24-29.

Kok, M. A., Chabot, N., & Lomber, S. G. (2014). Cross-modal reorganization of cortical afferents to dorsal auditory cortex following early-and late-onset deafness. *Journal of Comparative Neurology, 522*(3), 654-675.

Kral, A., Schröder, J.-H., Klinke, R., & Engel, A. (2003). Absence of cross-modal reorganization in the primary auditory cortex of congenitally deaf cats. *Experimental Brain Research, 153*, 605-613.

Lakatos, P., Chen, C.-M., O'Connell, M. N., Mills, A., & Schroeder, C. E. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron, 53*(2), 279-292.

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science, 320*(5872), 110-113.

Lakatos, P., Musacchia, G., O'Connel, M. N., Falchier, A. Y., Javitt, D. C., & Schroeder, C. E. (2013). The spectrotemporal filter mechanism of auditory selective attention. *Neuron, 77*(4), 750-761.

Lomber, S. G., Meredith, M. A., & Kral, A. (2010). Cross-modal plasticity in specific auditory cortices underlies visual compensations in the deaf. *Nature Neuroscience, 13*(11), 1421-1427.

Luo, H., Liu, Z., & Poeppel, D. (2010). Auditory cortex tracks both auditory and visual stimulus dynamics using low-frequency neuronal phase modulation. *PLoS Biology, 8*(8), e1000445.

Marchant, J. L., Ruff, C. C., & Driver, J. (2012). Audiovisual synchrony enhances BOLD responses in a brain network including multisensory STS while also enhancing target-detection performance for both modalities. *Human Brain Mapping, 33*(5), 1212-1224.

Meredith, M. A., & Allman, B. L. (2012). Early hearing-impairment results in crossmodal reorganization of ferret core auditory cortex. *Neural Plasticity, 2012*.

Meredith, M. A., & Allman, B. L. (2015). Single-unit analysis of somatosensory processing in the core auditory cortex of hearing ferrets. *European Journal of Neuroscience, 41*(5), 686-698.

Meredith, M. A., Clemo, H. R., Corley, S. B., Chabot, N., & Lomber, S. G. (2016). Cortical and thalamic connectivity of the auditory anterior ectosylvian cortex of early-deaf cats: Implications for neural mechanisms of crossmodal plasticity. *Hearing Research, 333*, 25-36.

Meredith, M. A., Kryklywy, J., McMillan, A. J., Malhotra, S., Lum-Tai, R., & Lomber, S. G. (2011). Crossmodal reorganization in the early deaf switches sensory, but not behavioral roles of auditory cortex. *Proceedings of the National Academy of Sciences, 108*(21), 8856-8861.

Meredith, M. A., & Lomber, S. G. (2011). Somatosensory and visual crossmodal plasticity in the anterior auditory field of early-deaf cats. *Hearing Research, 280*(1-2), 38-47.

Meredith, M. A., Nemitz, J. W., & Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *Journal of Neuroscience, 7*(10), 3215-3229.

Merrikhi, Y., Kok, M. A., Carrasco, A., Meredith, M. A., & Lomber, S. G. (2022). Multisensory responses in a belt region of the dorsal auditory cortical pathway. *European Journal of Neuroscience, 55*(2), 589-610.

Miller, L. M., & D'esposito, M. (2005). Perceptual fusion and stimulus coincidence in the cross-modal integration of speech. *Journal of Neuroscience, 25*(25), 5884-5893.

Navarra, J., Vatakis, A., Zampini, M., Soto-Faraco, S., Humphreys, W., & Spence, C. (2005). Exposure to asynchronous audiovisual speech extends the temporal window for audiovisual integration. *Cognitive Brain Research, 25*(2), 499-507.

Noesselt, T., Rieger, J. W., Schoenfeld, M. A., Kanowski, M., Hinrichs, H., Heinze, H.-J., & Driver, J. (2007). Audiovisual temporal correspondence modulates human multisensory

superior temporal sulcus plus primary sensory cortices. *Journal of Neuroscience, 27*(42), 11431-11441.

O'Connell, M. N., Barczak, A., Ross, D., McGinnis, T., Schroeder, C. E., & Lakatos, P. (2015). Multi-scale entrainment of coupled neuronal oscillations in primary auditory cortex. *Frontiers in Human Neuroscience, 9*, 655.

Pekkola, J., Ojanen, V., Autti, T., Jääskeläinen, I. P., Möttönen, R., Tarkiainen, A., & Sams, M. (2005). Primary auditory cortex activation by visual speech: an fMRI study at 3 T. *Neuroreport, 16*(2), 125-128.

Pienkowski, M., Shaw, G., & Eggermont, J. J. (2009). Wiener-Volterra characterization of neurons in primary auditory cortex using poisson-distributed impulse train inputs. *Journal of Neurophysiology, 101*(6), 3031-3041.

Powers, A. R., Hillock, A. R., & Wallace, M. T. (2009). Perceptual training narrows the temporal window of multisensory binding. *Journal of Neuroscience, 29*(39), 12265-12274.

Raposo, D., Sheppard, J. P., Schrater, P. R., & Churchland, A. K. (2012). Multisensory decision-making in rats and humans. *Journal of Neuroscience, 32*(11), 3726-3735.

Rebillard, G., Carlier, E., Rebillard, M., & Pujol, R. (1977). Enhancement of visual responses on the primary auditory cortex of the cat after an early destruction of cochlear receptors. *Brain Research, 129*(1), 162-164.

Romei, V., Gross, J., & Thut, G. (2012). Sounds reset rhythms of visual cortex and corresponding human visual perception. *Current Biology, 22*(9), 807-813.

Ross, L. A., Molholm, S., Butler, J. S., Del Bene, V. A., & Foxe, J. J. (2022). Neural correlates of multisensory enhancement in audiovisual narrative speech perception: a fMRI investigation. *Neuroimage, 263*, 119598.

Schormans, A. L., Scott, K. E., Vo, A. M., Tyker, A., Typlt, M., Stolzberg, D., & Allman, B. L. (2017). Audiovisual temporal processing and synchrony perception in the rat. *Frontiers in Behavioral Neuroscience, 10*, 246.

Stevenson, R. A., Altieri, N. A., Kim, S., Pisoni, D. B., & James, T. W. (2010). Neural processing of asynchronous audiovisual speech perception. *Neuroimage, 49*(4), 3308-3318.

Stevenson, R. A., & Wallace, M. T. (2013). Multisensory temporal integration: task and stimulus dependencies. *Experimental Brain Research, 227*, 249-261.

Stevenson, R. A., Zemtsov, R. K., & Wallace, M. T. (2012). Individual differences in the multisensory temporal binding window predict susceptibility to audiovisual illusions. *Journal of Experimental Psychology: Human Perception and Performance, 38*(6), 1517.

Stewart, D. L., & Starr, A. (1970). Absence of visually influenced cells in auditory cortex of normal and congenitally deaf cats. *Experimental Neurology, 28*(3), 525-528.

van Atteveldt, N. M., Formisano, E., Blomert, L., & Goebel, R. (2007). The effect of temporal asynchrony on the multisensory integration of letters and speech sounds. *Cerebral Cortex, 17*(4), 962-974.

Vroomen, J., Keetels, M., De Gelder, B., & Bertelson, P. (2004). Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Cognitive Brain Research, 22*(1), 32-35.

Wallace, M. T., & Stevenson, R. A. (2014). The construct of the multisensory temporal binding window and its dysregulation in developmental disabilities. *Neuropsychologia, 64*, 105-123.

Williams, P. E., Mechler, F., Gordon, J., Shapley, R., & Hawken, M. J. (2004). Entrainment to video displays in primary visual cortex of macaque and humans. *Journal of Neuroscience, 24*(38), 8278-8288.

Wong, C., Chabot, N., Kok, M. A., & Lomber, S. G. (2015). Amplified somatosensory and visual cortical projections to a core auditory area, the anterior auditory field, following early-and late-onset deafness. *Journal of Comparative Neurology, 523*(13), 1925-1947.

Zhou, Y., Carney, L. H., & Colburn, H. S. (2005). A model for interaural time difference sensitivity in the medial superior olive: interaction of excitatory and inhibitory synaptic inputs, channel dynamics, and cellular morphology. *Journal of Neuroscience, 25*(12), 3046-3058.

3.8. Supplementary Information

Supplementary Table 1. Electrode counts with RMS differences between after-flip clicks and before-flip baselines for each of five response windows

Subject	Response Windows	RMS attenuation			RMS enhancement		
		$p < .001$	$p < .01$	$p < .05$	$p < .001$	$p < .01$	$p < .05$
1	1	5	**7**	**9**	2	3	**6**
	2	4	**10**	**11**	**12**	**14**	**22**
	3	2	**6**	**10**	**6**	**8**	**11**
	4	**6**	**8**	**12**	1	5	**7**
	5	1	5	**7**	2	5	**11**
2	1	0	0	0	4	**13**	**22**
	2	0	1	1	**18**	**25**	**31**
	3	0	0	0	**19**	**29**	**31**
	4	1	1	1	**9**	**10**	**17**
	5	1	3	4	**7**	**8**	**15**

Bold, Fisher's exact test showed significant difference from 0 ($p < .05$).

Supplementary Table 2. Descriptive statistics for delay groups

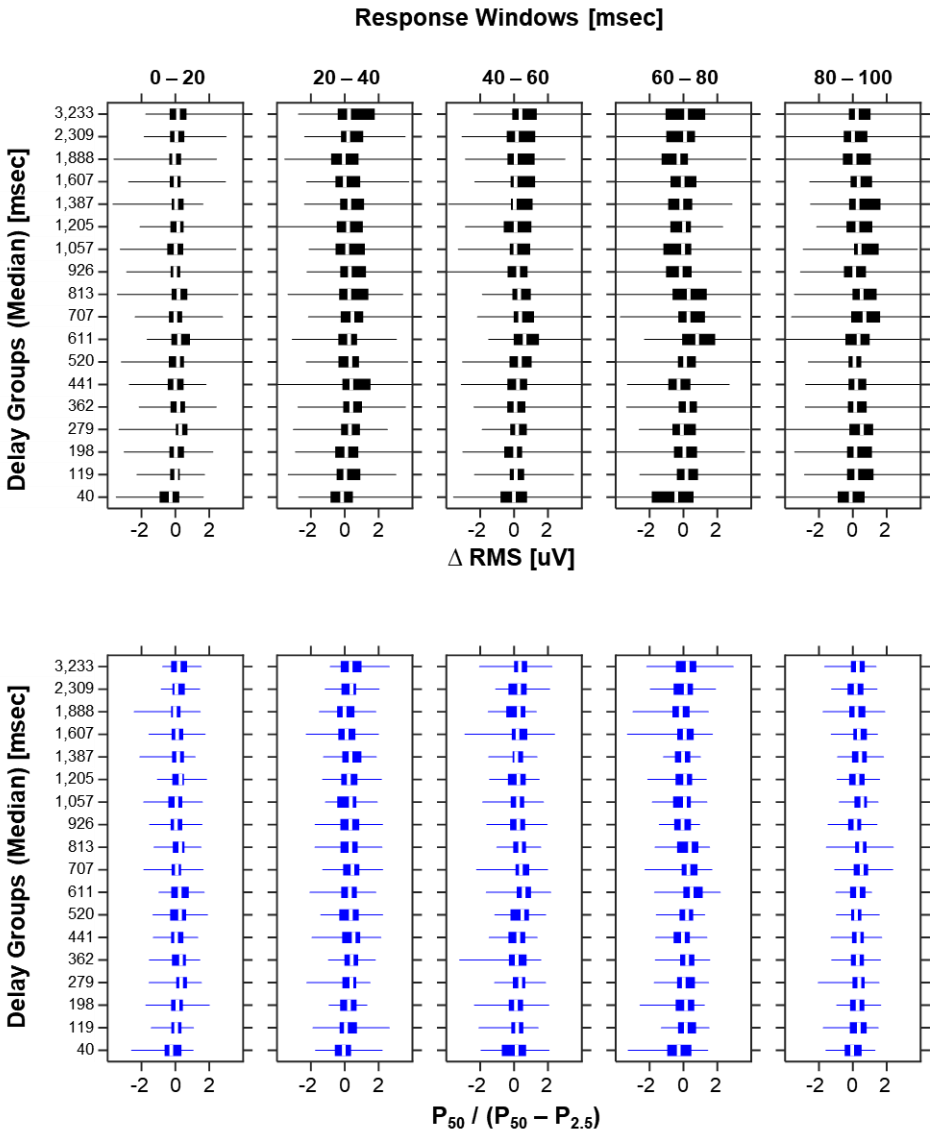
Subject	Delay Group	Full Range (msec)		Median (msec)	Inter-quartile Range (msec)
Jaime	1	0.0	78.3	39.9	37.6
	2	78.4	158.8	119.3	39.6
	3	159.1	238.1	197.5	39.5
	4	238.1	320.9	279.0	39.8
	5	320.9	402.0	361.8	38.6
	6	402.1	478.7	440.5	39.7
	7	478.8	563.0	520.1	42.4
	8	563.1	657.5	610.6	46.5
	9	657.5	759.7	707.2	53.2
	10	759.7	870.4	813.3	57.4
	11	870.7	987.2	925.8	54.8
	12	987.3	1124.5	1057.1	69.7
	13	1124.9	1294.7	1205.4	78.1
	14	1294.8	1491.5	1387.4	100.1
	15	1491.5	1747.0	1607.3	126.0
	16	1747.2	2074.4	1888.1	157.5
	17	2074.6	2664.4	2309.2	282.4
	18	2664.8	5693.3	3233.0	901.4
Snow	1	0.3	76.2	37.5	37.3
	2	76.3	152.3	113.2	38.0
	3	152.4	234.3	192.0	40.9
	4	234.4	312.0	274.2	37.4
	5	312.0	396.9	353.1	41.2
	6	396.9	483.0	439.5	44.2
	7	483.1	573.6	525.4	44.9
	8	573.8	672.8	620.9	52.3
	9	672.9	782.0	727.3	57.6
	10	782.0	894.7	837.9	56.1
	11	894.8	1032.7	962.3	67.1
	12	1032.8	1183.2	1103.6	77.6
	13	1183.2	1360.2	1269.2	87.7
	14	1360.4	1573.0	1462.4	103.7
	15	1573.1	1846.9	1697.4	132.6
	16	1847.3	2218.7	2023.1	183.9
	17	2219.3	2831.3	2490.1	318.8
	18	2831.9	4912.5	3346.8	687.9

Supplementary Table 3. Maximal electrode counts with RMS differences across delay groups for each of five response windows.

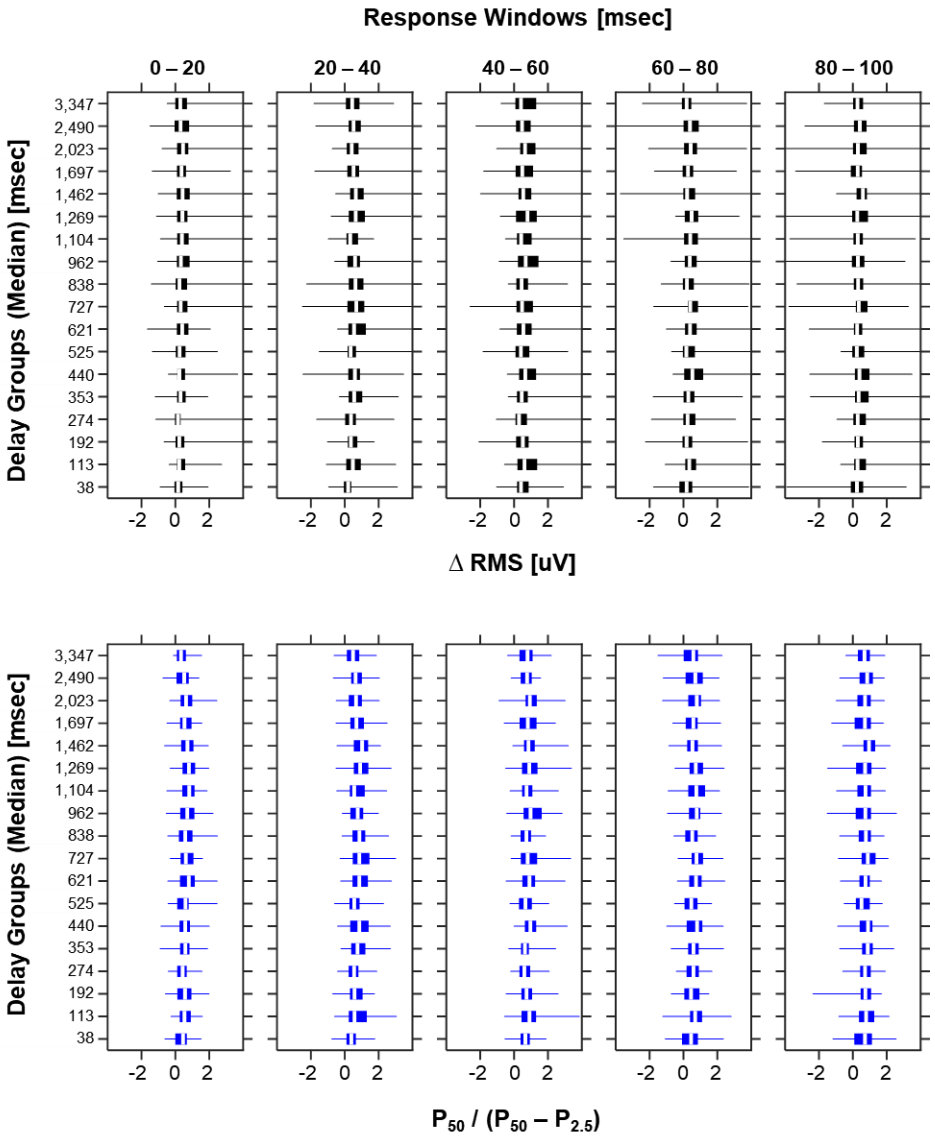
Subject	Response Windows	RMS decrease			RMS increase		
		$p < .001$	$p < .01$	$p < .05$	$p < .001$	$p < .01$	$p < .05$
1	1	2	**6**	**8**	1	2	**10**
	2	1	1	**7**	2	**7**	**15**
	3	3	**6**	**11**	3	5	**18**
	4	**6**	**9**	**15**	5	**9**	**19**
	5	1	1	2	2	5	**12**
2	1	0	0	0	2	3	**25**
	2	0	0	0	**7**	**15**	**32**
	3	0	0	0	**13**	**22**	**32**
	4	0	1	3	4	**10**	**23**
	5	0	0	2	**9**	**13**	**30**

Bold, Fisher's exact test showed significant difference from 0 ($p < .05$).

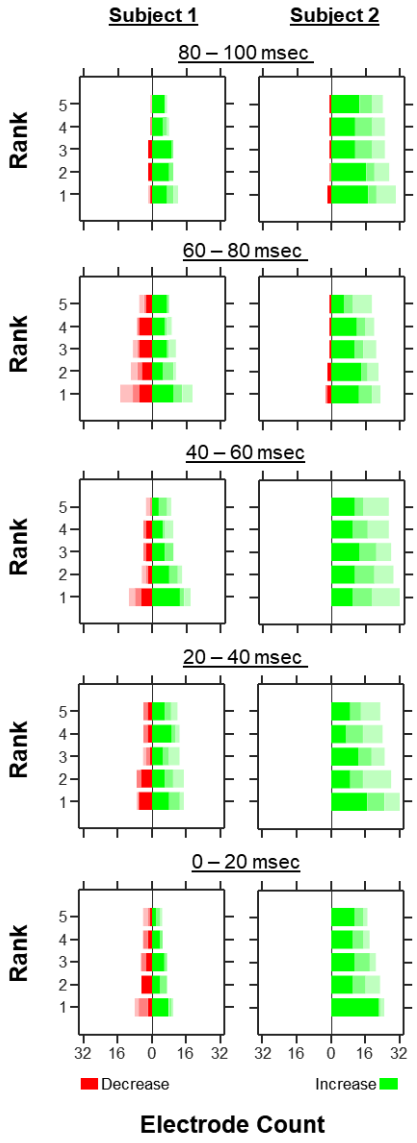
Underlined: count greater than the same subject and the same response window in Supplementary Material 1.



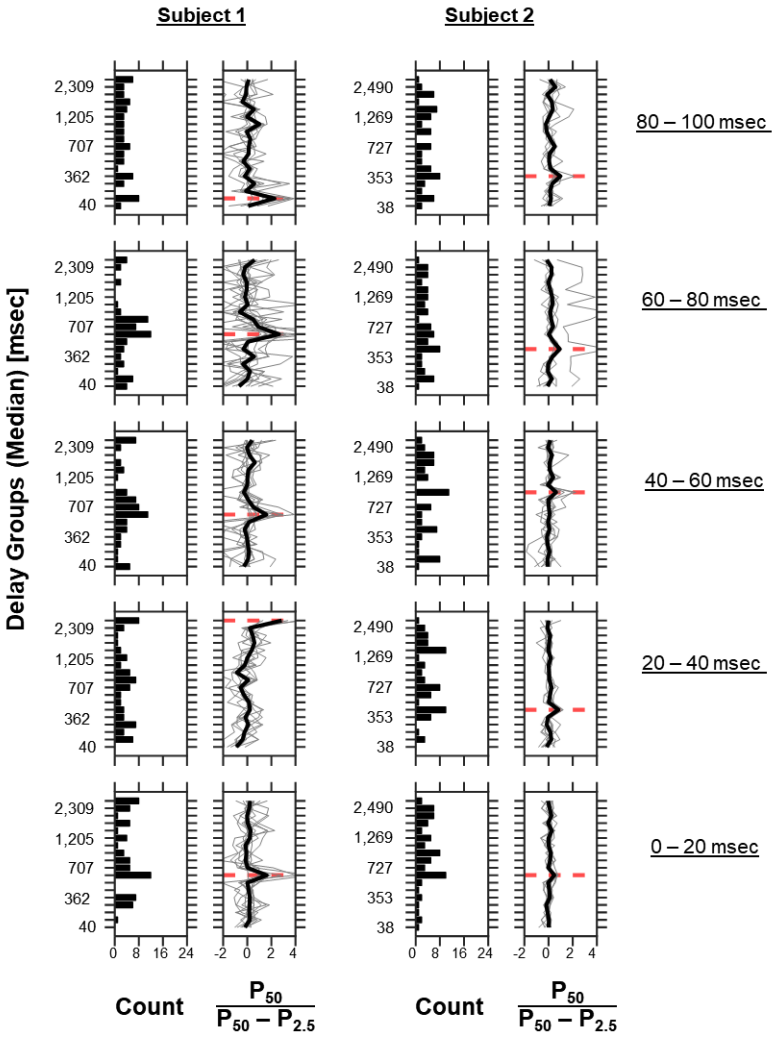
SI Figure 1. Effect of visual modulation in ΔRMS by flash-to-click delay groups in Subject 1. For conventions, see Figure 2C.



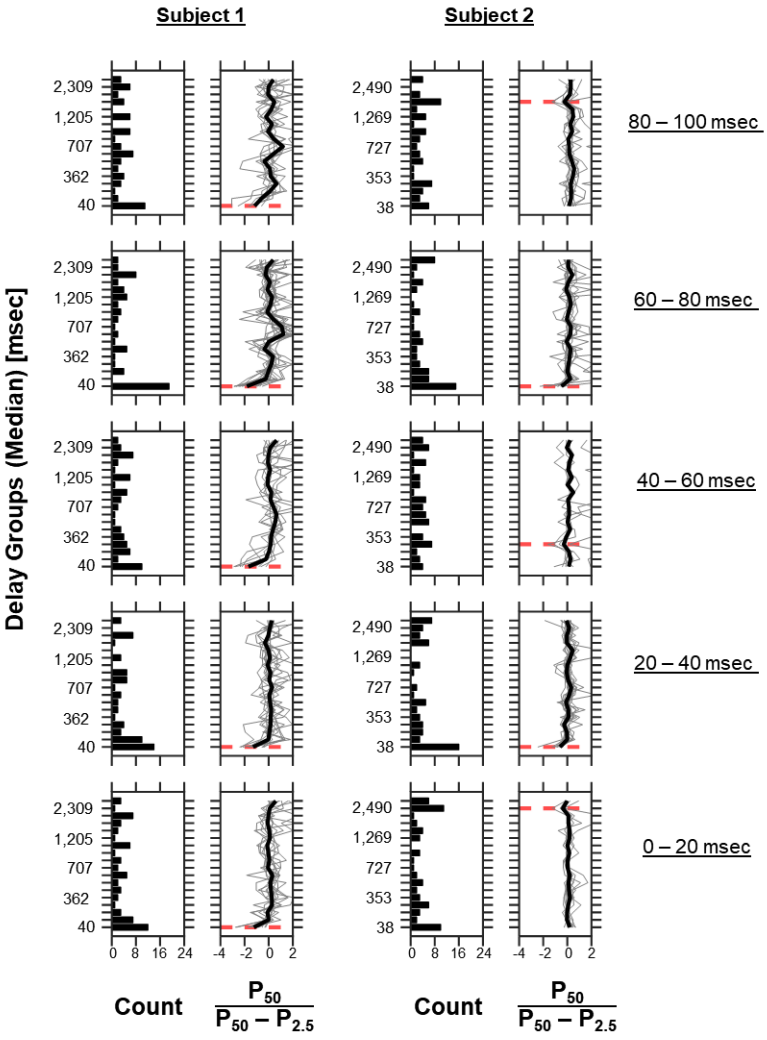
SI Figure 2. Effect of visual modulation in ΔRMS by flash-to-click delay groups in Subject 2. For conventions, see Figure 2C.



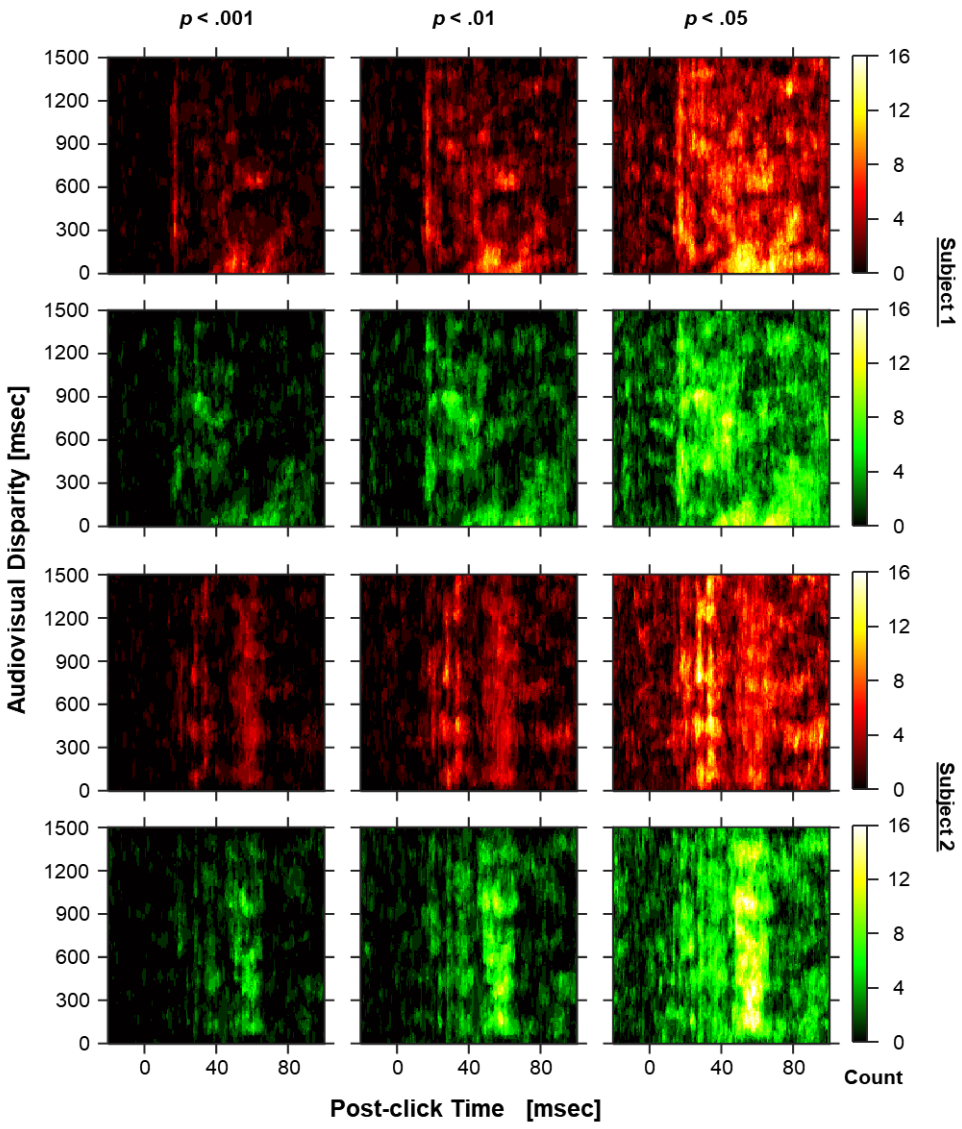
SI Figure 3. Effect of specifying audiovisual disparity on the detection of visual modulation. The top 5 counts of electrode with ΔRMS beyond inter-trial consistency thresholds for each of five response windows in Subject 1 (left) and Subject 2 (right). Three different shadings from dark to bright represented for 0.1%, 1% and 5% thresholds (i.e., the alpha values).



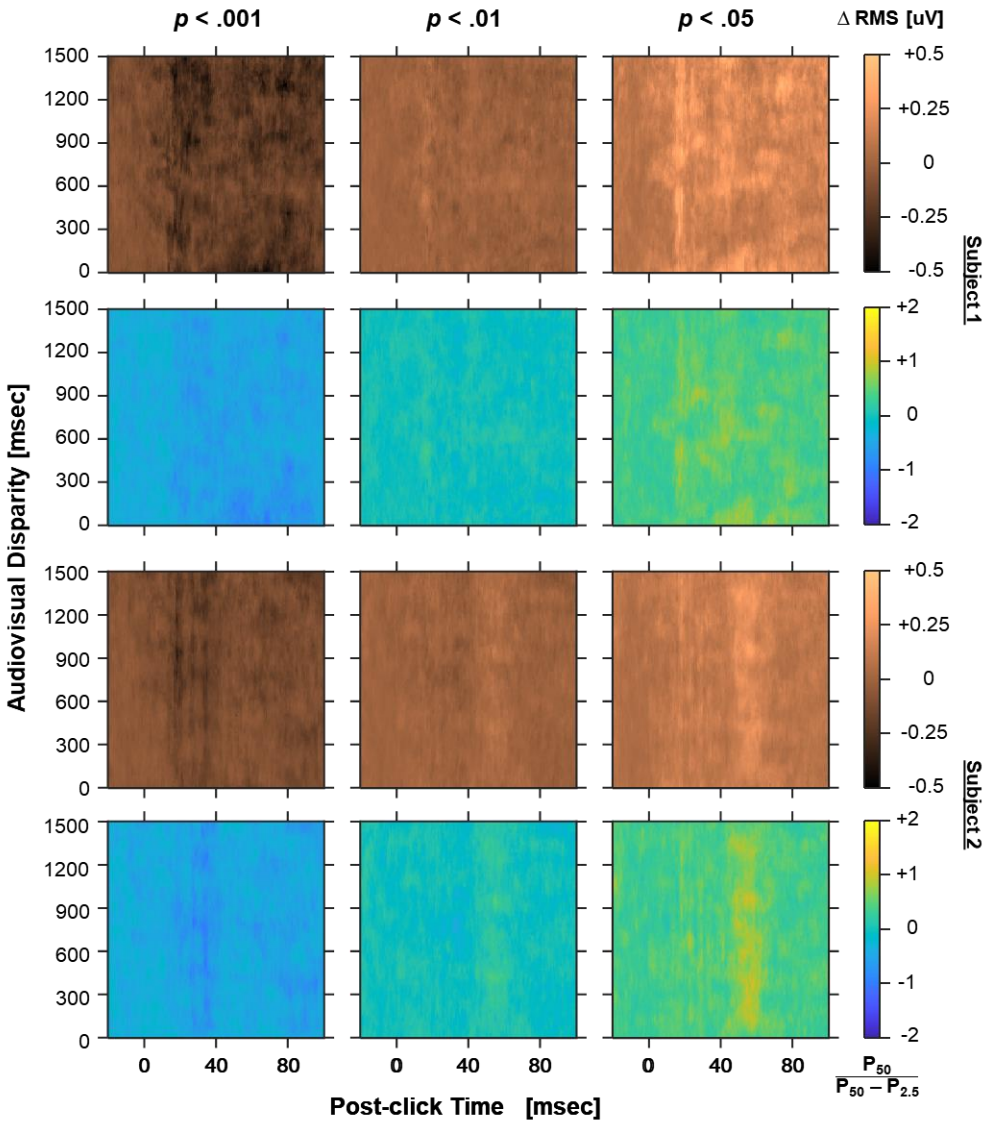
SI Figure 4. Distribution of optimal delays for RMS enhancement in each response window. For conventions, see Figure 4 legends.



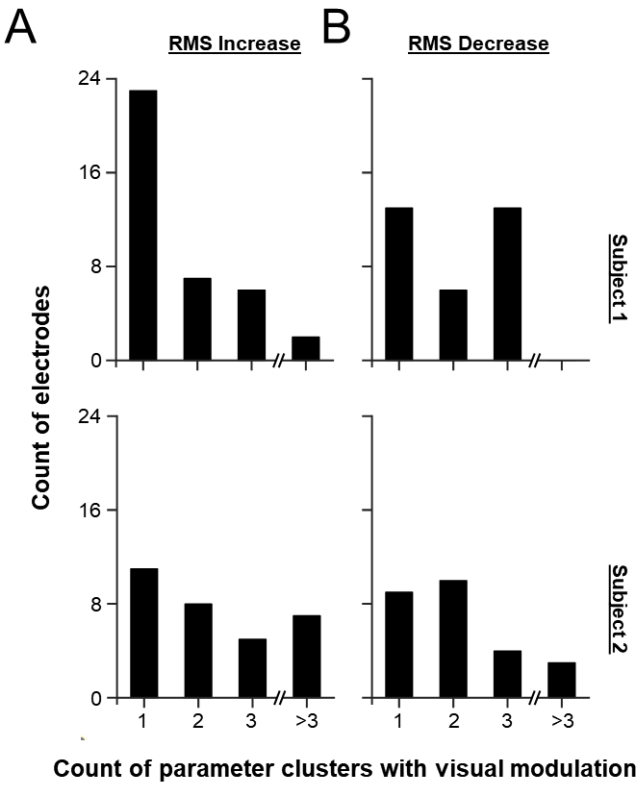
SI Figure 5. Distribution of optimal delays for RMS attenuation in each response window. For conventions, see Figure 4 legends.



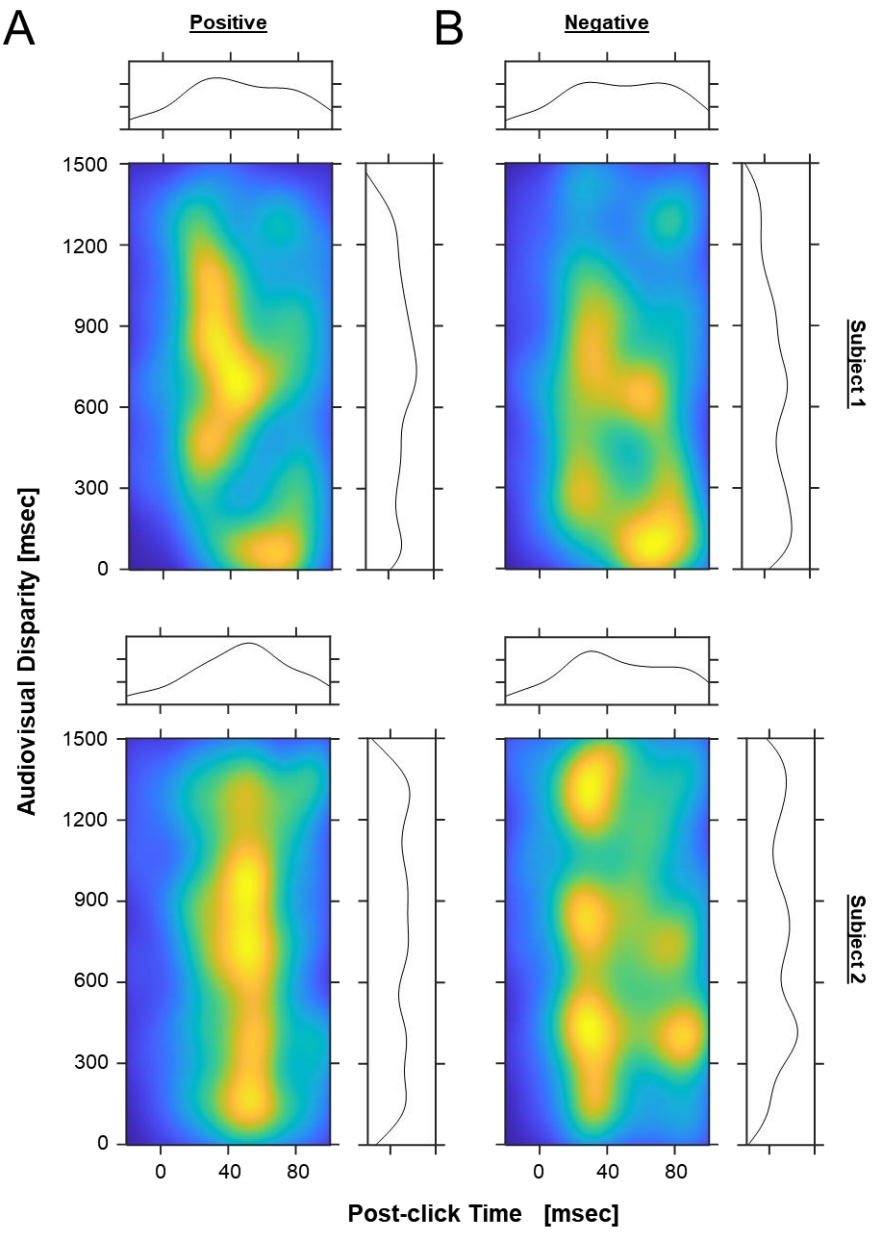
SI Figure 6. Estimated visual modulation as a function of audiovisual disparity with different threshold of inter-trial consistency. The count of electrodes with visual modulation beyond inter-trial consistency thresholds for each time point in both subjects. Three different consistency thresholds, 0.1% (left), 1% (middle), and 5% (right) were used. For conventions, see Figure 6.



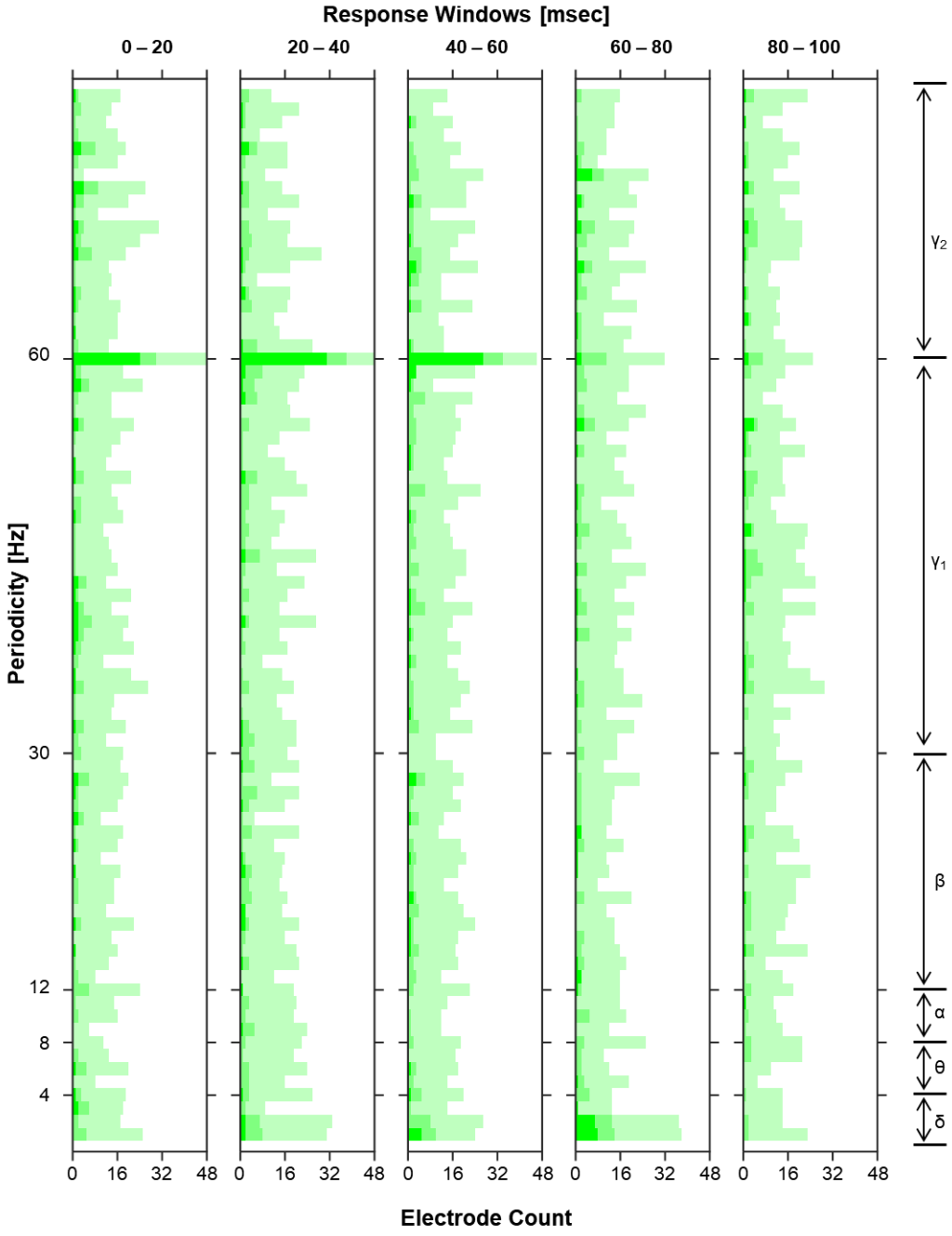
SI Figure 7. Population data of the estimated visual modulation as a function of audiovisual temporal disparity. LFP differences and their effect-size indices for each time point in both subjects. Three different consistency thresholds, 0.1% (left), 1% (middle), and 5% (right) were used. LFPs estimated for different audiovisual disparities on the vertical axis were compared to the same before-flip baseline.



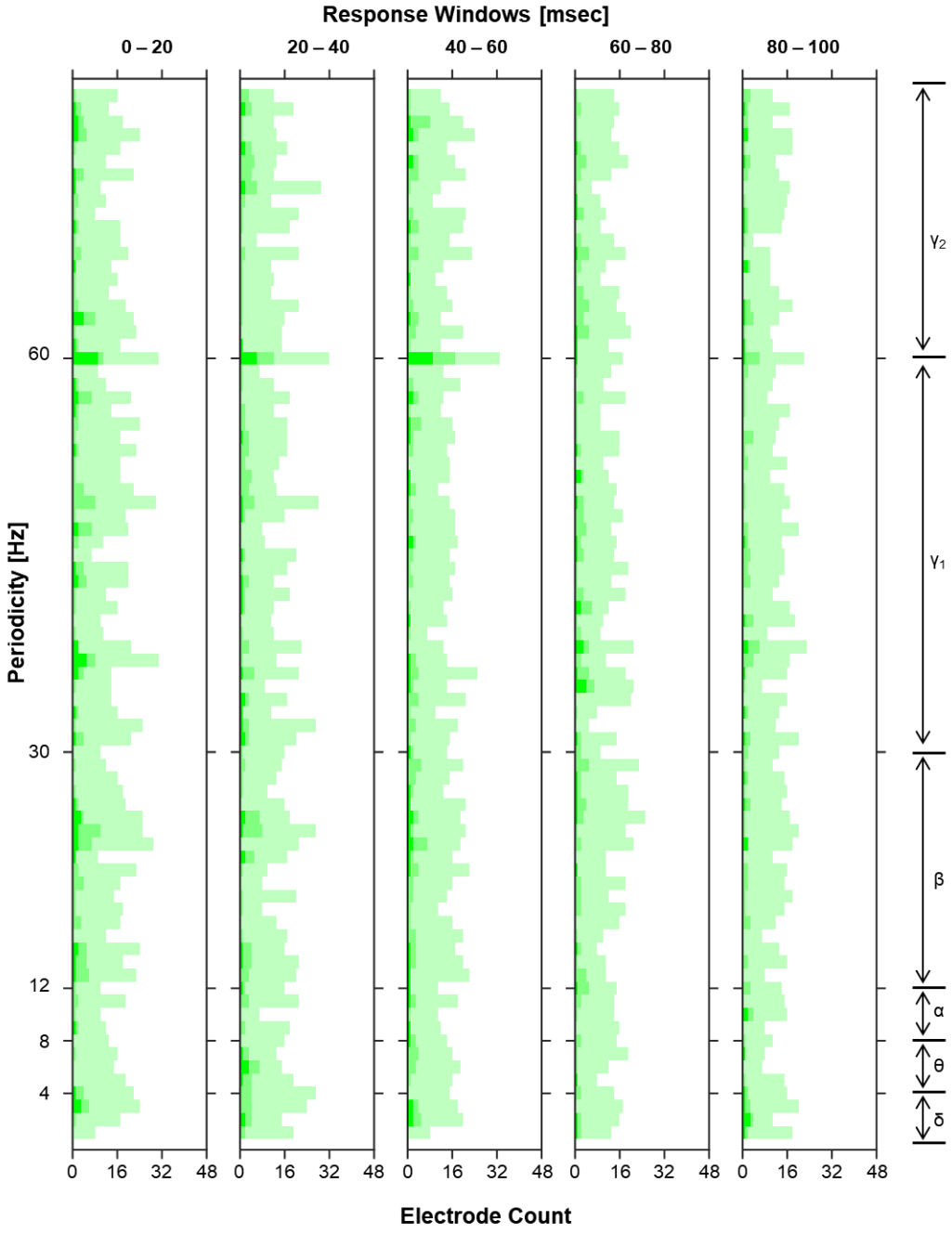
SI Figure 8. Occurrence of multiple local optimal combinations of audiovisual disparity and responses latency for visual modulation. The count of electrodes in each subject with single and multiple optimal parameter clusters, which is defined as an area of at least 500-ms*ms size in the two-dimensional space coordinated by response latency and audiovisual disparity. Effective visual modulation is defined as positive (A) or negative (B) LFP changes beyond 0.1% inter-trial consistency threshold.



SI Figure 9. Kernel-estimated probability density function of optimal audiovisual disparity and response latency for visual modulation. Probability density functions (PDFs) for the peaks of parameter clusters in each subject. Two-dimensional PDFs were estimated for positive changes (A) and negative changes (B) in LFPs and color-coded. Top panel shows the one-dimensional kernel-estimated PDFs for response latency. Right panel shows the one-dimensional kernel-estimated PDFs for audiovisual disparity.



SI Figure 10. Frequency coupling of visual modulation between 1 and 80 Hz in Subject 1. For conventions, see Figure 7 legends.



SI Figure 11. Frequency coupling of visual modulation between 1 and 80 Hz in Subject 2. For conventions, see Figure 7 legends.

4. Chapter 4: Effect of isoflurane on auditory and visually evoked potentials in cat under dexmedetomidine sedation

4.1. Relation to the thesis

To record electroencephalogramic (EEG) signals in animal subjects, it is possible to use subdermal needle electrode (i.e., scalp-recorded EEG) as an alternative to skull-implanted intracranial EEG, which avoids the risk of invasive surgical procedures such as craniotomy. However, the animal subject will still need light sedation/anesthesia to allow the placement of needle electrodes (analgesia) and to ensure reliable auditory and visual stimulation during recording (akinesia). The effect of anesthetics has been a common and inevitable confounding factor in many animal studies and has received much attention from neuroscientists. Meanwhile, anesthesia is a research topic by itself, and its mechanism is closely related to that of consciousness and sleep. This chapter, in general, is aimed at characterizing the effect of a common anesthetics (i.e., isoflurane) during the recording of auditory and visually evoked potentials.

Specifically, this chapter, as a part of my thesis, is also a pilot study for verifying whether isoflurane should be used for studying visual modulation of auditory evoked potentials. Before investigating the audiovisual processing using evoked potentials in cats, it is necessary to first quantify the effect anesthetic agents on different measurements of evoked potentials. Isoflurane has been long used as an anesthetic agent in research animals due to the ease of use, while more recently it has become more and more common in veterinary care to use sedative agents, such as dexmedetomidine, for less invasive procedures. Although both isoflurane and dexmedetomidine can provide the animal subjects with the status needed for the recording procedure, their anesthetic effects as a confounding factor and the limitations on experiment designs are drastically different.

Isoflurane is known to act mainly on GABAergic receptors, but also glutamatergic receptors, and potassium channels throughout the entire brain, although its exact mechanism of action is still unclear. It is administrated through a mask as an inhalable anesthetic, and therefore can provide a more stable but deeper anesthesia when compared to dexmedetomidine. On the other hand, dexmedetomidine is an alpha-2 agonist, and its sedative effect was found to necessarily involve a brainstem nucleus called locus coeruleus. With an intramuscular injection of dexmedetomidine, the animal subject experienced sedation more like natural sleep. Due to this feature, the use of dexmedetomidine is often mandatory as pre-medication before anesthesia induction.

To help the choice of anesthetic agents for studying multisensory processing, the effect of isoflurane on both auditory evoked potentials and visual evoked potentials were studied in cats under dexmedetomidine sedation. Although the aim of this chapter deviates from the major research question of the thesis, the outcome of this project will provide knowledge on selections of anesthetics for future neuroscience studies of this kind, explore the potential use of multi-modal evoked potentials in intraoperative monitoring in clinical settings, and most importantly, make an informed decision on which agent to use in Chapter 5.

This manuscript has been submitted with revision to *Frontiers in System Neuroscience* and is currently under review.

4.2. Abstract

Evoked potentials can be used as an intraoperative monitoring measure in neurological surgery. Auditory evoked potentials (AEPs), or specifically brainstem auditory evoked responses

(BAERs), are known for being minimally affected by anesthetics, while visually evoked potentials (VEPs) are presumed to be unreliable and easily affected by anesthetics. While many anesthesia trials or intraoperative recordings have provided evidence in support of these hypotheses, the comparisons were always made between AEPs and VEPs recorded sequentially rather than simultaneously. Although simultaneous recording of AEPs and VEPs may be a challenge in clinical settings, it is completely feasible in animal models.

Five cats under dexmedetomidine sedation received five, 10-min blocks of isoflurane with varying concentrations (0.5%, 1.0%, 1.5%, 2.0%, and 0% as in recovery) while click-evoked AEPs and flash-evoked VEPs were recorded simultaneously from subdermal electrodes. Unlike checkerboard flipping as in Chapter 3, an impulse of flash stimulates visual system diffusively, and therefore was more suitable in subjects under sedation. We found that, in terms of their waveforms, (1) short-latency AEPs (BAERs) was the least affected while middle-latency AEPs was dramatically altered by isoflurane, and (2) short-latency VEPs was less persistent than that of short-latency AEPs, while both middle- and long-latency VEPs were largely suppressed by isoflurane and, in some cases, completely diminished. In addition, the signal strength in all but the middle-latency AEPs was significantly suppressed by isoflurane. We identified multiple AEP or VEP peak components demonstrating suppressed amplitudes and/or changed latencies by isoflurane. Overall, we confirmed that, as in human, cat VEPs are overall more affected during isoflurane anesthesia than short-latency AEPs (i.e., BAERs).

4.3. Introduction

For the intraoperative monitoring of anesthesia, brainstem auditory evoked potentials (BAEPs) and somatosensory evoked potentials (SSEPs) are commonly used and are preferred over

visual evoked potentials (VEPs) (Banoub, Tetzlaff, & Schubert, 2003; Sloan & Heyer, 2002). Over three decades ago, it was demonstrated that the amplitude of VEPs and cortical SSEPs (Sebel, Flynn, & Ingram, 1984; Sebel, Ingram, Flynn, Rutherfoord, & Rogers, 1986) but not BAEPs (Sebel et al., 1986) are decreased during surgical level anesthesia. The same effect of anesthetics on VEP amplitudes was also shown by more recent studies (Chi & Field, 1990; Tanaka et al., 2020; Tenenbein, Lam, Klein, & Lee, 2006; Wiedemayer, Fauser, Armbruster, Gasser, & Stolke, 2003). The susceptibility of VEPs to anesthetics has brought challenges to some surgical procedures where the monitoring of intact visual function is essential, such as surgery to remove a cancerous tumor near the optic nerve. Moreover, the confounding effect of anesthetics has long been a persistent concern for neuroscientists studying sensory functions in anesthetized animals, and has motivated many investigations on this issue (Grasshoff & Antkowiak, 2006; Jehle et al., 2009; Santangelo et al., 2018; Villeneuve and Casanova, 2003).

Unlike the auditory or somatosensory systems, most retinal inputs bypass subthalamic processing. In macaques, before visual inputs reach primary visual cortex (V1), the vast majority of retinal ganglion neurons are subcortically relayed in the lateral geniculate nucleus (LGN) with a minority (less than 10%) projecting to superior colliculus (Perry & Cowey, 1984). It is now generally recognized that sub-cortical nuclei along ascending sensory pathways are less affected by anesthetics, or that neurons distant from the cortex are less affected by anesthetics. While halothane and enflurane suppress cortically generated SSEPs measured from central C3 electrode with reference electrode at forehead, they did no suppress cervically generated SSEPs from spinal C2 electrode (Samra, Vanderzant, Domer, & Sackellares, 1987; Sebel, Erwin, & Neville, 1987). Similarly in the auditory system, the amplitude of auditory middle latency responses (MLRs), which is identified to be generated by the auditory thalamus and primary cortical areas, is

decreased by propofol or isoflurane (Schwender, Kaiser, Klasing, Peter, & Poppel, 1994); this was not the case for BAEPs (Manninen, Lam, & Nicholas, 1985; Sebel et al., 1987).

One of the technical features that are particular to AEP recordings is that AEPs originating from subthalamic, thalamic, and cortical sources can be recorded using the same electrode configuration (Caron-Desrochers, Schönwiesner, Focke, & Lehmann, 2018). This electrode configuration allows VEP recordings as well. Taking this feature as an advantage, it is possible to compare simultaneously-recorded BAEPs, MLRs, cAEPs. The use of simultaneous recording minimizes the confounding factors or variations introduced by different subjects, different electrodes, and different levels of anesthesia.

In this study, we anesthetized five cats under four different isoflurane concentrations and performed 10 EEG recording sessions for each concentration. One recording session comprised a 57-second train of asynchronized click and flash pulses (**Fig. 1**). EEG signals were filtered off-line using three different band-pass filters to acquire VEPs and AEPs in a wide range of latency window. This stimulus paradigm allowed us to compare the dynamics of isoflurane effect along cat auditory and visual pathway as reflected in EEG signal with minimized confounding variables. We found that both AEPs and VEPs were affected by isoflurane anesthesia. While VEP magnitude was overall suppressed, AEP magnitude was preserved but with the waveforms largely altered.

4.4. Methods

All procedures were conducted in compliance with the National Research Council's Guide for the Care and Use of Laboratory Animals (8th edition; 2011) and the Canadian Council on Animal

Care's Guide to the Care and Use of Experimental Animals (1993. Furthermore, the following procedures were also approved by Animal Care Committee (DOWB) for the Faculty of Medicine and Health Sciences at McGill University.

Animal preparation and anesthesia protocol

Five cats of 3-to-5 years old, including four females, were used in this study. After subject was sedated by 0.04mg/kg dexmedetomidine (Dexdomitor, Zoetis) injected intramuscularly, the left eye was occluded using a black contact lens so that visual stimuli were presented unilaterally. Phenylephrine (Mydfrin, Alcon) was applied to the right eye to dilate the pupil, and saline drops were used as lubrication.

Subjects were placed on a water-circulated heating pad (TP-400, Gaymar). Once vital signs (heart rate and SpO2) were stable, we carried out two 15-minute recording blocks while the subject was breathing 80% oxygen (Dispomed) as a baseline under dexmedetomidine, i.e., Dex condition. Next, the oxygen was mixed with isoflurane (AErrane, Baxter) at four inspired concentrations (0.5%, 1%, 1.5%, and 2%) using an isoflurane vaporizer (Isotec 4, Smiths Medical). A 10-minute recording block was carried out for each concentration (except that for subject No. 1 each recording block was 15 minutes). Finally, another 10-minute block was recorded after the termination of isoflurane, i.e., "0%" condition, to investigate the effect of anesthesia recovery (**Fig. 1A**). Subject's vital signs and electrode impedance were checked at the end of each recording block. Each recording block started immediately after changing the concentration setting on the vaporizer. At the end of data collection, electrodes and contact lens were removed before the atipamezole (Antisedan, Zoetis) was administrated intramuscularly to facilitate recovery from dexmedetomidine sedation.

Visual and auditory stimuli

The visual stimuli consisted of flashes presented to subjects from a 5-mm-diameter light-emitting diode (~11 degrees of visual field, LED, DigiKey). The intensity of flash stimuli was calibrated to 10 lux by adjusting the magnitude of a .3-ms-long squared pulse as the input signal to the LED. The auditory stimuli were 300-us-long clicks emitted by an 8-cm-diamter loudspeaker (Fostex). The sound level of the click stimuli was calibrated to 55 dB SPL using a sound meter (Model 2250, B&K). Both auditory and visual stimulation signals were generated by the same processor (RZ2, TDT). The LED was attached to the top of the loudspeaker and placed at 8-cm away from subject at the direction of 45 degrees right to the midline.

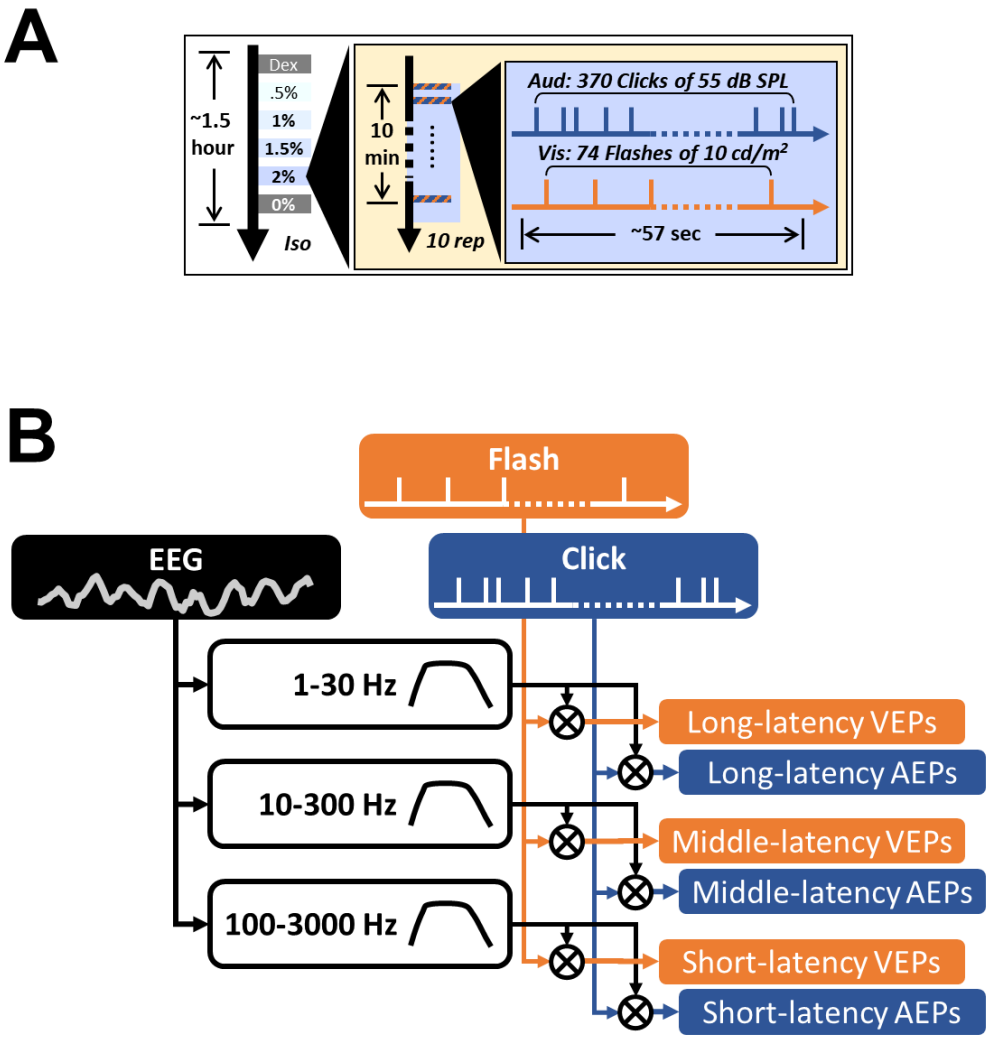


Figure 1. Experiment Design and Data Analysis. (A) The timeline of the auditory and the visual stimulus trains used in the current study. The three concentric boxes contain the timeline of each session, each block, and the entire recording in each subject, respectively. Dex, dexmedetomidine only. See the detailed breakdown of the recording time in the Supplementary Figure 7. **(B)** An illustration of the data analysis pipeline on deriving the AEPs and the VEPs. The cross inside a circle indicates cross-correlation. All filled boxes are signals. The unfilled boxes indicate digital filters with varying pass-bands.

To manipulate the timing of auditory and visual stimulus, two independent, 57-second-long pulse trains for triggering clicks and flashes, respectively, were pre-made in Matlab using a Poisson random process and loaded into the stimulus/recording software (Synapse, TDT). The auditory stimulus train contained 370 click pulses and the visual stimulus train contained 74 flash pulses. The inter-pulse deadtime was set to 20 msec and 300 msec for the auditory and visual stimuli, respectively. The auditory and the visual stimulus trains always started and stopped

simultaneously in each session for the current dataset. The exact same session was repeated 10 times under each isoflurane conditions (Dex, 0.5%, 1%, 1.5%, 2%, 0%, in order) at a rate of ~1 session per minute (**Fig. 1A**). The four isoflurane concentrations correspond to 0.31, 0.63, 0.95, and 1.27 times of minimum alveolar concentration (MAC) in cats.

EEG recording and signal processing

Stainless steel needles (25G) were placed subdermal as recording electrodes. The active electrode was placed near the midpoint of subject's interaural line, while the reference electrode was placed beneath the right ear (ipsilateral to the side of visual stimulation). The ground electrode was placed on subject's dorsum (~10 cm behind should blade near the midline). The impedance of both active and reference electrodes was maintained below 3 kOhm during recording. The signal was amplified and digitized with a pre-amplifier (TDT, Medusa4Z) at ~6.1 kHz, streamed onto a digital signal processor (TDT, RZ2), and stored on a computer hard drive.

All data analysis was performed offline. Signal was digitally notched at 30 Hz, 60 Hz, 120 Hz, 180 Hz, and 240 Hz before passing through three different band-pass filters (1-30 Hz, 10-300 Hz, and 100-3000 Hz) for long-, middle-, and short-latency responses, respectively (**Fig. 1B**). For short-latency responses, the original signal was upsampled to ~24.4 kHz before filtering. Epochs of various windows were extracted surrounding either click or flash onsets, and averaged to derive auditory or visually evoked potentials.

Data analysis

Root-mean-square values were obtained using MATLAB built-in function rms(). Each averaged waveform was separated into a pre-stimulus zero-mean baseline window and a post-stimulus response window, producing two RMS values respectively. The corrected RMS value was calculated as in the following equation:

$$RMS_{corrected} = \sqrt{RMS_{response}{}^2 - RMS_{baseline}{}^2}$$

Noted that although the mean of the pre-stimulus baseline window was zero due to the baseline correction, its RMS was not, and therefore can be used as an estimation of stimulus-unrelated signal power. This baseline power could be due to the spontaneous neural activities or thermal noise.

To analyze individual components of response waveforms, we first manually determined a window surrounding the candidate peak or trough of our interest by interactively overlaying a pair of vertical cursors on the averaged waveform. This process was carried out by the first author and was assisted with visual guidance from graphical information including polarity, magnitude, timing in the context of the entire waveform and other components previously identified. Bounded by this window, we identified the component representatives (e.g. peak, or trough) by looking for the smallest absolute value of the first order derivative using MATLAB built-in function min(). A component representative derived in this way can be either a local extreme, (i.e., zero-crossing first order derivative), or a local plateau or less-abrupt slope, in which case one component blends into another.

In order to identify peak component semi-automatically, we first determined the peak component of the grand averaged waveform across subjects. Then, the peak time of a given subject was determined by the nearest local maxima to the peak time landmark derived from the grand average, which was also used as a new landmark in determining the peak time of each block for the given subject.

For characterizing minute-by-minute change, curve fitting of peak amplitudes as a function of minutes after isoflurane termination was performed in MATLAB and its Curve Fitting Toolbox™. We first obtained the initial values for the model coefficients using function polyfit(), where peak amplitude values were first converted into logarithmic space for exponential modeling. Then, function fit() was used to fine tune the model coefficients with customized anonymous functions $Amp = a \cdot e^{b \cdot Time} + c$ for exponential modeling. A quadratic model $AMP = a \cdot Time^2 + b \cdot Time + c$ was also used as a secondary approach.

Statistics

Due to a small sample size of subjects (N = 5) in this study, the Friedman test was chosen to examine the statistical significance of difference among isoflurane blocks, since each subject was repeatedly measured for all six different isoflurane blocks. In each subject, we also examined the effect of isoflurane blocks on peak time, amplitude, and RMS values derived from session averages. Since ten sessions per block were measured independently, the Kruskal–Wallis test was used. Both tests are non-parametric versions of one-way analysis of variance (ANOVA), and are available in Statistics and Machine Learning Toolbox™ on MATLAB as friedman() and kruskalwallis(). Corresponding to either test, multiple comparison tests were carried out using Tukey's honestly significant difference procedure using MATLAB function multcompare().

Finally, comparisons of curve fitting coefficients were examined by Wilcoxon signed rank test using MATLAB function signrank(). P-values smaller than 0.05 were reported as statistically significant.

4.5. Results

Using the anesthesia protocol and stimulus configuration described in the methods section, we recorded auditory and visually evoked potentials with subdermal needle electrodes in five feline subjects. Through three different band-pass filters, short-, middle-, long-latency AEPs and VEPs were obtained from EEG signals. To evaluate the effect of isoflurane, we will first compare signal strength quantified from the entire waveform among blocks, before going through each individual peak component that were selected. While the narrative in this study mainly focuses on the population level using the block-averaged waveforms from individual subjects, the session-averaged waveforms were also analyzed to reveal the trend of isoflurane effects in each subject that may be either the same or opposite to the trend from the population.

Effect of isoflurane on waveforms and signal strength

With EEG signals band-pass filtered between 1 Hz and 30 Hz, we observed long-latency AEPs (LL-AEPs, or elsewhere referred as cAEPs) in all five subjects. The waveforms of LL-AEPs were viewed in the window from 100-msec before to 500-msec after click onsets, and were highly consistent across subjects. Each LL-AEP waveform is characterized by a main positive peak component with a peak time around 20-ms (P20) followed by a less prominent but observable

negative component around 80-ms (N80). Later peak components after 100-ms latency were either completely absent or present but with a negligible magnitude that we could not quantify with its signal-to-noise ratio.

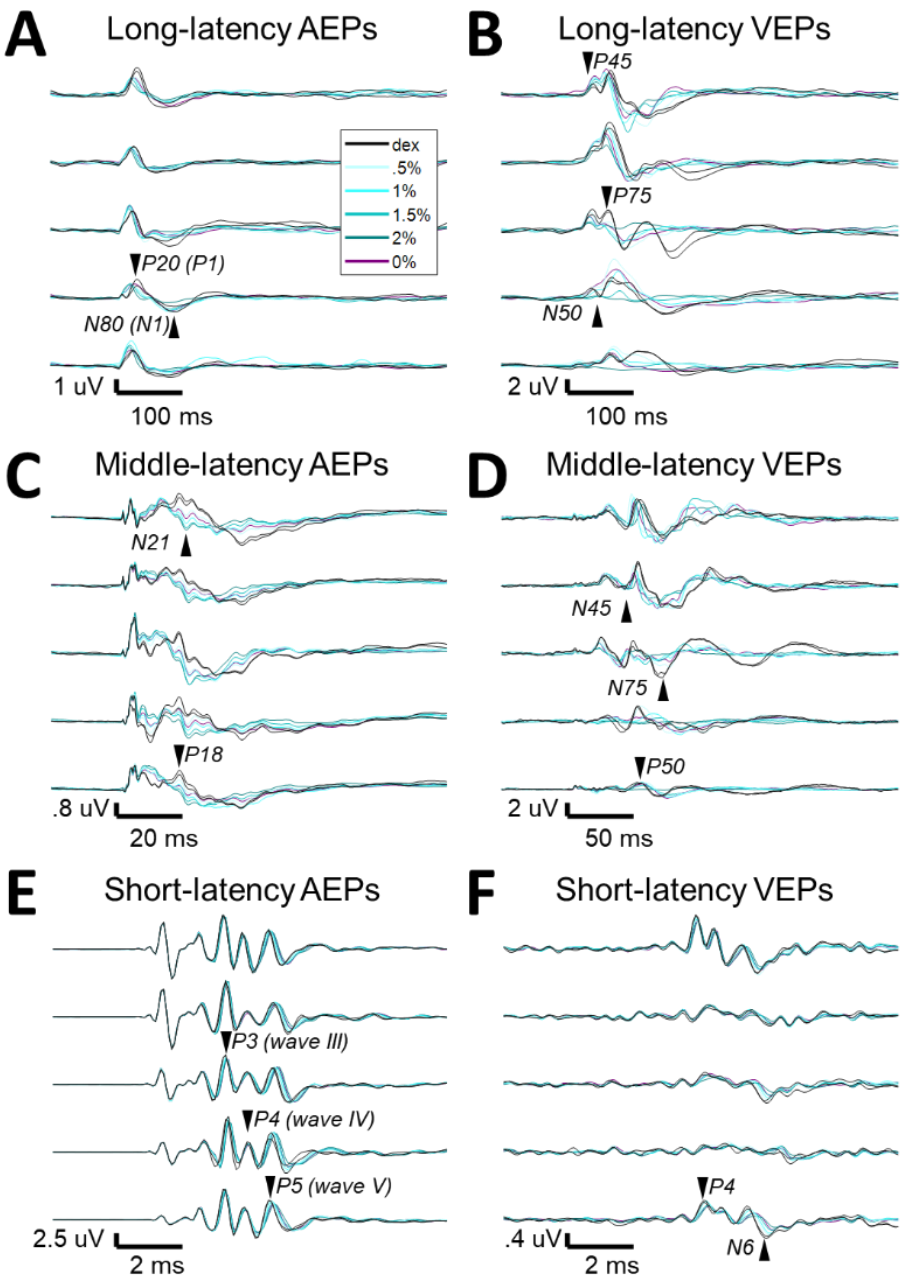


Figure 2. The waveforms of AEPs and VEPs from three different filters. (A-F) Each of the five rows demonstrates the waveforms of AEPs or VEPs derived from one of the three filters for one subject, from bottom to top. Noted that different time and amplitude scales were used. The isoflurane/anesthesia treatment is color-coded. Black, baseline block with only dexmedetomidine (i.e., Dex) administrated. Blue with light-to-dark shading, the isoflurane blocks with up-stepping concentrations. Purple, the block after isoflurane stops. Arrows and labels indicate the selected peak components for further analysis.

The administration of isoflurane did not abolish LL-AEPs (**Fig. 2A**), but consistently decreased the peak time of P20. The overall signal strength was also affected in both directions depending on both isoflurane concentrations and subjects. Consequently, the effect of isoflurane on Root Mean Square (RMS) values, which we used to quantify signal strength of LL-AEPs, was small but significant ($Q = 13.57$, $p < .05$). Multiple comparison test showed that the difference in only one pair of comparison, that is the 2%- versus the 1%-isoflurane block, had statistical significance ($p < .05$) (**Fig. 3A Right**).

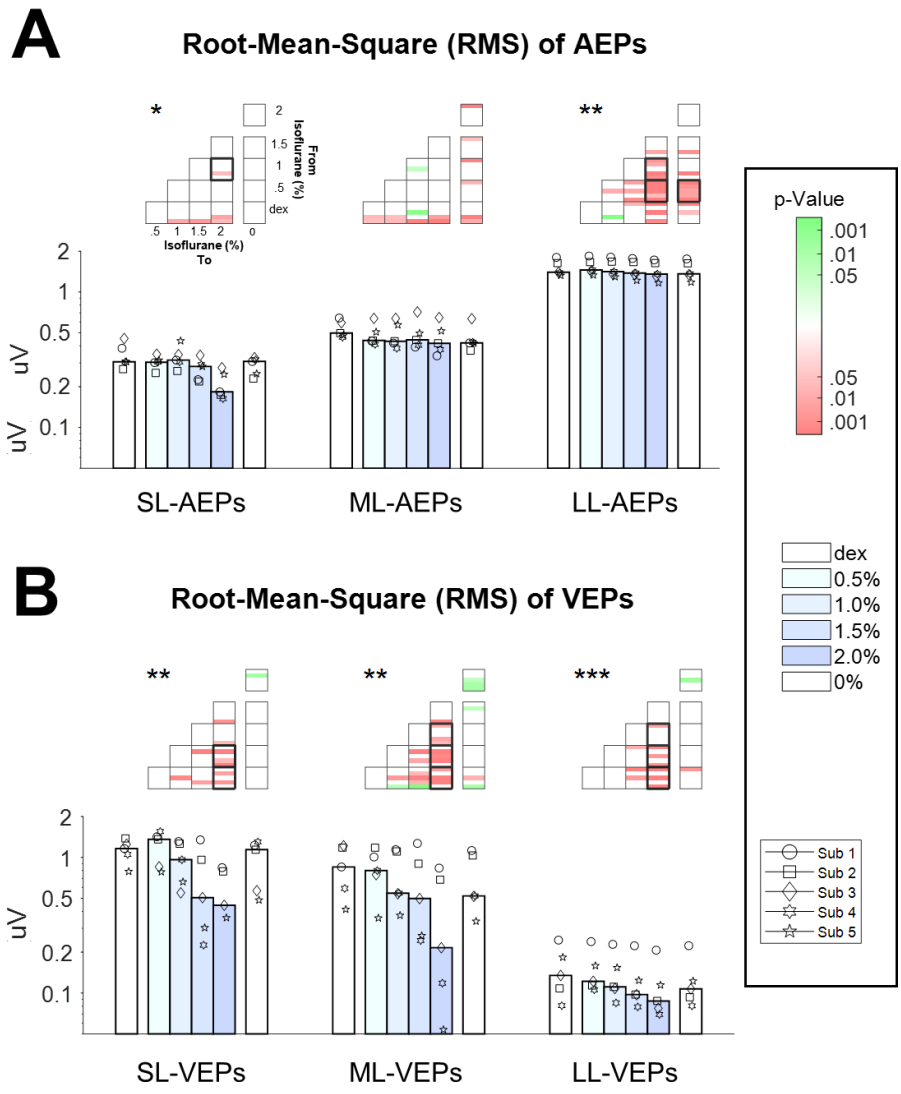


Figure 3. Effect of isoflurane on the signal strength of AEPs and VEPs. The bottom axis shows the block median of the RMS values across subjects from 6 blocks filled with different shading of blue indicating isoflurane concentration. Dex, dexmedetomidine. From the left to right, the data from the three filters were clustered for AEPs (A) and VEPs (B). Dots with different shapes, data of individual subject. The heat maps on the top show the P-values from multiple comparison tests for individual subject, between the conditions labeled on the right and at the bottom. Green, increase. Red, decrease. Black border, a significant pair in the multiple comparison tests from the group statistics ($p < .05$). Both ordinates are in logarithmic scale.

Using the same filter and window parameters as LL-AEPs, we observed long-latency VEPs (referred as LL-VEPs here) in all five subjects as well. Unlike LL-AEPs, the waveforms of LL-VEPs were extremely inconsistent across subjects. In four out of five subjects, the waveform of LL-AEPs shared a common pattern, which is featured by two positive components with peak implicit time about 45-msec and 75-msec, i.e., P45 and P75, and a negative component between the two, i.e., N50. Despite the comparable peak times, the relative amplitudes of these three peak components still varied largely between subjects.

The effect of isoflurane on LL-VEPs was overall suppressive but dependent on isoflurane concentration, and was inconsistent across subjects (**Fig. 2B**), and the inter-subject variance could be contributed by many sources. For example, there could have individual variation in the placement of recording electrodes, anatomical positions of the primary and the secondary visual cortex. These variance suggested that normalization for each subject, which was not employed here, can be a useful approach for quantifying the effect of isoflurane. The statistical significance, however, was examined using repeated-measure non-parametric methods.

The main block effect on RMS values of LL-VEPs was statistically significant ($Q = 16.89$, $p < .01$) (**Fig. 3B Right**). From both the waveforms and the RMS values, only the 2%-isoflurane block showed substantial attenuation of LL-VEPs across all five subjects. In subject No. 4, LL-VEP was completely absent during the 2%-isoflurane block. Multiple comparison test showed that there was only significant decrease in RMS values during 2%-isoflurane block when compared to dexmedetomidine (i.e., Dex) block ($p < .05$) or 0.5%-isoflurane block ($p < .01$).

With EEG signals band-pass filtered between 10 Hz and 300 Hz, we observed middle-latency AEPs (ML-AEPs or elsewhere referred as MLAEPs and MLAERs) in all five subjects.

The waveforms of ML-AEPs were viewed in the window between 20-msec before and 100-msec after click onsets. The waveforms of ML-AEPs did not resemble those recorded from human, but were very consistent across subjects. Each ML-AEP waveform was composed of three phases. After click onsets, the first phase, between 0- and 10-msec, was characterized by a positive sharp spike corresponding to BAEP wave V, followed by a smaller positive peak component showing more variance across subjects compared to the previous one. The second phase, from 10- to 30-msec, is characterized by three pairs of peak and troughs (i.e., P13 and N16, P18 and N21, and P23 and N26), which appeared to be the most affected phase out of the three. The last phase was a long shallow negative component with peak time larger than 30-msec.

The administration of isoflurane dramatically altered the ML-AEP waveforms consistently across subjects (**Fig. 2C**) without changing the overall signal strength (**Fig. 3A**). In the window between 5-ms and 10-ms after click onsets, which comprised the later component in the first phase, the response amplitude shifted upwards and became more positive. Oppositely, in the window between 10-ms and 30-ms after click onsets, the response amplitude shifted downwards and became more negative. However, the pattern of peak components as described above were preserved to a large extent. It is worth noting that the overall signal strength was also not affected despite of the alteration of the waveform. There was no statistically significant effect of isoflurane block on the RMS values of ML-AEPs (**Fig. 3A Middle**).

Using the same filter as ML-AEPs but a longer epoch window from 50-msec before to 250-msec after flash onsets, we observed a second version of VEPs, which we termed as middle-latency VEPs (ML-VEPs). Unlike ML-AEPs, the waveforms of ML-VEPs had much less consistency across subjects, but slightly more than LL-VEPs. The most identifiable peak components (N45,

P50, and N75) observed in ML-VEPs could also be found in LL-VEPs but with their polarities inverted.

The effect of isoflurane on ML-VEPs was quite similar to LL-VEPs (**Fig. 2D**) but had a few noticeable differences. Same as in LL-VEPs, the RMS values of ML-VEPs were overall significantly decreased ($Q = 15.86$, $p < .01$) (**Fig. 3B Middle**). In the 2%-isoflurane block, ML-VEP was almost abolished in subject No. 5, but not in subject No. 4. In the 1.5%-isoflurane block, while subject No. 2, No. 3, and No. 4 demonstrated significant decreases in RMS values, subject No. 1 demonstrated a significant increase in RMS values. Multiple comparison test showed that RMS values in the 2%-isoflurane block was significantly lower than Dex ($p < .01$), 0.5%- ($p < .05$) and 1%-isoflurane blocks ($p < .05$).

Finally, with EEG signals band-pass filtered between 100 Hz and 3000 Hz, we observed short-latency AEPs (SL-AEPs or commonly known as ABRs, BAERs, or BAEPs) using the epoch window from 2-msec before to 10-msec after click onsets. The waveforms of SL-AEPs in all five subjects had highly consistent pattern, which was composed of five distinct pairs of positive and negative peak components (i.e., wave I to V). Their peak times were comparable with those in the literature (Fullerton, Levine, Hosford-Dunn, & Kiang, 1987).

The effect of isoflurane on SL-AEP waveforms was small but prominent and consistent across subjects (**Fig. 2E**). The later components (e.g., wave IV and V) were delayed and prolonged, while the earlier components seemed minimally affected. Although the RMS values of SL-AEPs were decreased for less than 7% on average, the effect of isoflurane block was statistically significant ($Q = 19.51$, $p < .01$) (**Fig. 3A Left**). Multiple comparison test showed that the decrease

found in the 2%-isoflurane block was significant when compared to the 0.5%- ($p < .01$) or the 1%-isoflurane blocks ($p < .05$).

Using the same filter and window parameters as SL-AEPs, we surprisingly observed evoked potentials time-locked to flash onsets, which we referred to as short-latency VEPs (SL-VEPs). The waveform of SL-VEPs was detectable in all five subjects, but their signal strength had a large variance across subjects. We identified a positive (P4) and a negative (N6) peak component around 4-msec and 6-msec after flash onsets, which were most apparent in subject No. 1 (**Fig. 2F bottom row**) and No. 5 (**Fig. 2F top row**) but still discernible in the other subjects. The pattern of the peak components was more consistent than ML- and LL-VEPs across all five subjects, except that the N6 component had a later peak time in subject No. 4.

The effect of isoflurane on this so-called SL-VEPs was comparable to SL-AEPs (**Fig. 2F**) but with a larger suppression. The later component N6 was delayed and prolonged by isoflurane more than the earlier component P4. Despite of being the lowest among all six evoked potentials, the RMS values of SL-AEPs was decreased significantly ($Q = 22.26$, $p < .001$) (**Fig. 3B Left**). Multiple comparison test showed that only those in the 2%-isoflurane block was significantly decreased from the Dex ($p < .01$), the 0.5%- ($p < .01$) and the 1%-isoflurane blocks ($p < .05$), although the decrease seemed to be negatively correlated to the isoflurane concentration.

Thus far, we found the waveform and signal strengths of the six AEPs/VEPs were all affected by isoflurane in some way. During the 2%-isoflurane block, the RMS values clearly decreased in all but ML-AEPs. RMS value is a quantification applied to the entire waveform and requires minimal arbitrary supervision except for the selection of epoch window. The change of RMS value does not indicate which peak component(s) are specifically affected during the

experiment. Therefore, in the following sections, we further examined the effect of isoflurane by selecting at least two peak components from each of the six AEPs/VEPs. For each peak component, we determined the implicit time of the peak (or peak time) for the averaged waveforms from each subgroup as the nearest local maxima or minima to the peak time determined from the average of all the subgroups (see **Method** section). The peak amplitude was defined as the waveform magnitude at the peak time.

Long-latency AEPs

We only quantified the first two peak components (i.e., the P20 and the N80, also known as P1 and N1) from LL-AEPs, as the later components were not consistently accessible. Only P20 peak time but not peak amplitude (**Supplementary Fig. 1A Left, 4B Left**) was significantly decreased by isoflurane ($Q = 24.54$, $p < .001$). Multiple comparison test showed a significant decrease in P20 peak time in the 2%-isoflurane block, when compared to the Dex block ($p < .001$) or the 0.5%-isoflurane blocks ($p < .01$), but not the 1%-isoflurane block (**Supplementary Fig. 1B Left**). This pattern was inversed from what we observed with the RMS values, in which case the decrease was only significant from the 2%- to the 1%-isoflurane blocks. Also, the decrease from the Dex to the 1.5%-isoflurane block was also significant ($p < .01$), suggesting that P20 peak time could be affected by isoflurane of a lower concentration than the overall RMS value.

We also examined the N80 component, and we found its amplitude (**Supplementary Fig. 1A Right**) was significantly suppressed by isoflurane ($Q = 18.14$, $p < .01$). Same as in P20 peak time, N80 amplitude was significantly lower in the 2%-isoflurane block when compared to the Dex ($p < .01$) or the 0.5%-isoflurane block ($p < .01$), suggesting that P20 peak time and N80 amplitude may have an innate conjugation, but not with the overall RMS value. There was also a

block effect on N80 peak time (**Supplementary Fig. 1B Right**) with a weak statistical significance ($Q = 12.86$, $p < .05$). However, multiple comparison test only revealed significant difference ($p < .01$) between the 0%- and the 2%-, and between the 0%- and the 1%-isoflurane block. Our findings from the P20 and the N80 components suggested that the small suppression in the signal strength of LL-AEPs may originate from the decrease in both N80 amplitudes and P20 peak time.

Long-latency VEPs

The amplitude of the P75 component from LL-VEPs was significantly attenuated by isoflurane in general, with a slight increase during the 0.5%-isoflurane block (**Supplementary Fig. 2A Right**). We found the block effect of isoflurane on P75 amplitude with medium statistical significance ($Q = 17.57$, $p < .01$). Multiple comparison test showed that a significant decrease occurred during the 2%-isoflurane block, when compared to the Dex ($p < .05$) or the 0.5%-isoflurane blocks ($p < .01$).

A similar pattern can be found in the P45 and the N50 components (**Supplementary Fig. 2A Left and Middle**), despite of the identification of these two were particularly difficult in subject No. 5 with a low signal-to-noise-ratio in its atypical waveform. Except for a slight increase during the 0.5%-isoflurane block, both P45 amplitude ($Q = 12.09$, $p < .05$) and N60 amplitude ($Q = 14.71$, $p < .05$) were overall attenuated by isoflurane. It is worth noting that multiple comparison tests revealed significant washout effect ($p < .05$) in the 0%-isoflurane block in comparison with the 2%-isoflurane block for all three peak amplitudes from LL-VEPs. Meanwhile, the peak time of the three showed no block effect of significance (**Supplementary Fig. 2B**). These findings suggested that the suppression of RMS values from LL-VEPs was contributed by all three components we selected.

Middle-latency AEPs

Whereas the RMS values of ML-AEPs were not suppressed by isoflurane, the waveforms of ML-AEPs were drastically altered. A numerous of peak components can be derived from ML-AEPs, but we only selected the amplitudes and peak times of the P18 and the N21 component for further quantitative analysis, as these two appeared to have the most representative change on waveform. Both peak amplitudes (**Supplementary Fig. 3A**) significantly shifted towards more negative values, that is, in the case of P18 amplitude, towards zero ($Q = 24.54$, $p < .001$), and, in the case of N21 amplitude, away from zero ($Q = 25.00$, $p < .001$). Multiple comparison tests showed that for both P18 and N21 amplitudes the difference were significant, between the 2%-isoflurane and the Dex blocks ($p < .001$), between the 2%- and the 0.5%-isoflurane blocks ($p < .01$), and between the 1.5%-isoflurane and the Dex blocks ($p < .01$). Additionally, P18 peak time (**Supplementary Fig. 3B Left**) demonstrated statistical significant decrease ($Q = 19.80$, $p < .01$). Multiple comparison test showed significant decrease between the 2%-isoflurane and the Dex blocks ($p < .01$), between the 2%- and the 0.5%-isoflurane blocks ($p < .05$), and between the 1.5%-isoflurane and the Dex blocks ($p < .01$). However, N21 peak time was unaffected (**Supplementary Fig. 3B Right**).

It is worth pointing out that the same positive findings observed as in the group statistics were followed by each of five subjects, with the only exception being subject No. 5 only for P18 peak time. Furthermore, the amount of peak amplitude shift appeared to be in linear relationship with the concentration of isoflurane, which was not often observed in the other components or the RMS values.

Taking the advantage of their small session-by-session variance, we also examined the recovery process (i.e., wash-out effect) of isoflurane on P18 and N21 peak amplitudes during the 0%-isoflurane block by examining data from each one-minute session. We found that both 10-point recovery functions may dictate the time course of isoflurane wash-out process (**Fig. 4**). After the termination of isoflurane administration (i.e., switching from 2% to 0% concentration), the recovery from anesthesia should follow an exponentially decay pattern while isoflurane was cleared out of the central neural system (Stoelting & Eger, 1969). As we expected, in this 10-minute recovery block, both P18 and N21 peak amplitudes recovered back towards the baseline level faster during the first half of the time than the second half. In some subjects, the recovery function approached to a plateau in the end.

We carried out curve fitting for the recovery functions using two nonlinear models (exponential function and second-order polynomial function), both of which have three coefficients. Both P18 and N21 peak amplitudes fit well with both two models, with adjusted R-squared values larger than 0.5 in all five subjects. However, considering that the exponential decaying pattern of the effective concentration is featured for common inhaled gas anesthetics, we only derived the coefficients from the exponential model for further comparison. Each of the three coefficients represents for gain, decay factor, and asymptotic value of the decay process. We found no statistically significant difference for any one of the three coefficients between the P18 and the N21 components.

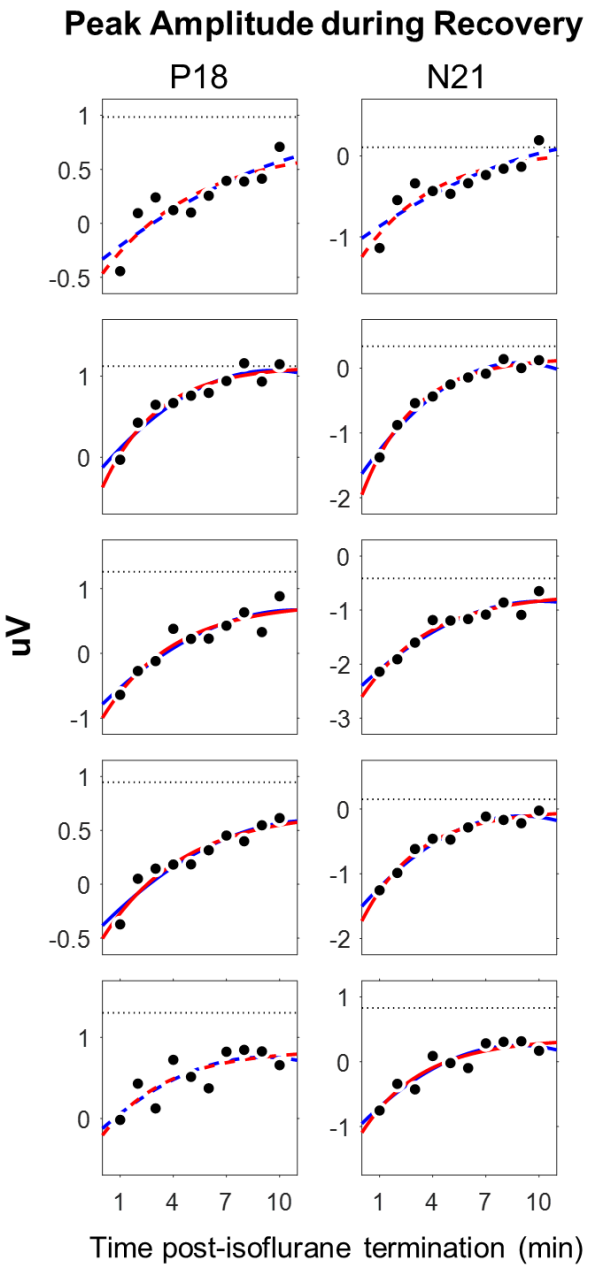


Figure 4. Time course of the wash-out effect in the ML-AEP components P18 and N21. Each row of panels corresponds to one subject. Dot, peak amplitude derived from session-averaged waveforms in the 0%-isoflurane block. Dotted line, data from the 0.5%-isoflurane block. Red, exponential mode. Blue, second order polynomial model. Solid, adjusted R^2 larger than 0.75. Dashed, adjusted R^2 between 0.5 and 0.75.

In summary, we found that the peak components of ML-AEPs, taking the P18 and the N21 as examples, could effectively quantify the change of ML-AEP waveforms caused by isoflurane, and may be particularly useful in predicting the anesthesia emergence.

Middle-latency VEPs

We quantified three peak components from ML-VEPs with peak times corresponding to those from LL-VEPs but inverted polarities. N75 amplitude (**Supplementary Fig. 4A Right**), like the P75 amplitude from LL-VEPs, was also significantly decreased ($Q = 12.54$, $p < .05$), but the following multiple comparison test showed statistical significance in the 2%-isoflurane block against the .5%- or the 1%-isoflurane block ($p < .05$), but not the Dex block. P50 amplitude (**Supplementary Fig. 4A Middle**) was also significantly decreased ($Q = 12.31$, $p < .05$), but instead of the recovery effect found in N50 amplitude from LL-VEPs, the following multiple comparison test revealed a significance decrease in P50 amplitude from the Dex to the 2%-isoflurane block ($p < .05$). Unlike the P45 amplitude from LL-VEPs, the effect of isoflurane on N45 amplitude (**Supplementary Fig. 4A Left**) was not statistically significant. Although none of the three peak times from LL-VEPs were affected by isoflurane, P50 peak time from ML-AEPs (**Supplementary Fig. 4B Middle**) was significantly increased ($Q = 12.70$, $p < .05$). Multiple comparison test showed a significant increase from the 0.5%- to the 2%-isoflurane blocks ($p < .05$) and a significant decrease from the 2%- to the 0%-isoflurane blocks ($p < .05$).

Together with what we found from LL-VEPs, these results suggested that the effect of isoflurane on VEPs could majorly reside in the spectral range that is overlapped between LL-AEPs and ML-AEPs, but the contribution from the non-overlapping spectral range may have also presented.

Short-latency AEPs

Among all five waves in SL-AEPs or BAEPs, wave III (P3) and V (P5) were most frequently reported in the literature. Here, we quantified P3 (wave III), P4 (wave IV), and P5 (wave V) amplitude and peak time for further analysis. Isoflurane significantly reduced P4 amplitude ($Q = 16.89$, $p < .01$) and P5 amplitudes ($Q = 17.11$, $p < .01$) but not P3 amplitudes (**Supplementary Fig. 5A**). Multiple comparison tests showed that only P5 amplitude was significantly lower in the 2%-isoflurane block when compared to the 0.5%- or the 1%-isoflurane block ($p < .05$). Decrease in P4 amplitude was observed in the 2%-isoflurane block but did not reach statistical significance. The effect of isoflurane on P4 amplitude was persistent after isoflurane was stopped during the 0%-isoflurane block when compared to the Dex ($p < .05$) or the 0.5%-isoflurane block ($p < .01$).

As for peak times, all of the three components (**Supplementary Fig. 5B**) were significantly affected, which includes P3 peak time ($Q = 21.21$, $p < .01$), P4 peak time ($Q = 13.52$, $p < .05$), and P5 peak time ($Q = 24.15$, $p < .001$). They all demonstrated a significant increase from the Dex to the 2%-isoflurane block, which includes P3 peak time ($p < .01$), P4 peak time ($p < .05$), and P5 peak time ($p < .001$), as revealed by multiple comparison tests. For the P3 and the P5 components, the increase in peak times from the Dex baseline level reached statistical significance in the 1.5%-isoflurane block ($p < .05$), which suggested that these two are likely more sensitive to isoflurane than P4 peak time. The increases were also persistent after isoflurane termination during the 0%-isoflurane block for both P3 peak time ($p < .01$) and P5 peak time ($p < .05$). P5 peak time also showed a significant difference between the 0.5%- and the 2%-isoflurane blocks ($p < .01$).

In short, we found the P5 component most sensitive to isoflurane in SL-AEPs, in terms of both peak amplitude and peak time. The other components, such as the P3 and the P4 components,

were affected by isoflurane as well. However, we don't consider the effect size observed in the current study, especially for peak amplitudes, was substantially larger than those reported previously. Rather, the statistically positive change may be due to low signal noise and hence stronger statistical power.

In previous studies, peak III-to-V latency was frequently used as a quantification for intraoperative ABR monitoring in surgery room. Our data showed that, not only both P3 and P5 peak times themselves were increased, P3-to-P5 latency was increased as well ($Q = 17.81$, $p < .01$) (**Fig. 5 Left**). Multiple comparison test showed a significant increase in P3-to-P5 latency in the 2%-isoflurane block when compared to the Dex ($p < .01$) and the 0.5%-isoflurane block ($p < .05$), which suggested that the increase in P5 peak time did not entirely derive from the increase in P3 peak time.

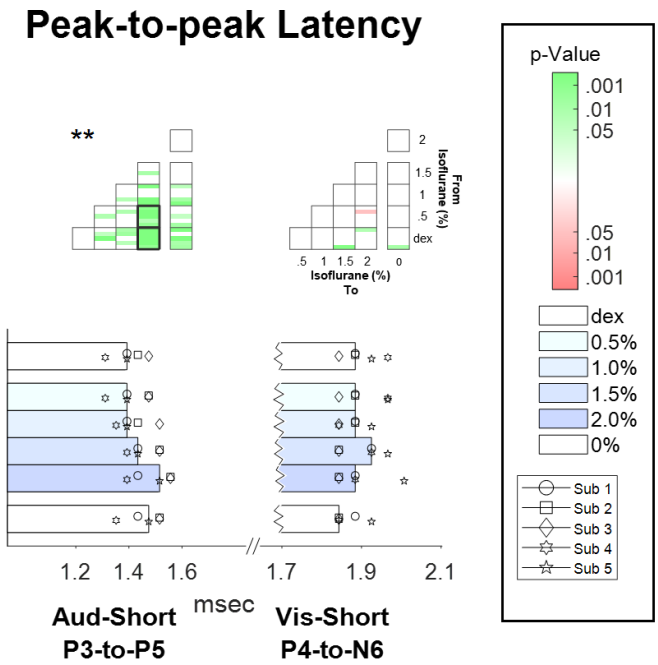


Figure 5. Effect of isoflurane on the peak-to-peak latencies in SL-AEPs and SL-VEPs. Same legends in Fig. 3 were used here.

Short-latency VEPs

We only quantified the P4 and the N6 components from SL-VEPs, as the other peak components cannot be consistently seen in all subjects. While P4 amplitude seemed to be spared by isoflurane, N6 amplitude (**Supplementary Fig. 6A Right**) was systematically attenuated by isoflurane ($Q = 22.94$, $p < .001$). Multiple comparison test revealed that the N6 amplitude in the 2%-isoflurane block was significantly lower than the Dex or the 0.5%-isoflurane blocks ($p < .01$).

Isoflurane also significantly increased both P4 ($Q = 22.90$, $p < .001$) and N6 ($Q = 22.63$, $p < .001$) peak times (**Supplementary Fig. 6B**). Compared to the Dex block, P4 peak time was significantly increased in the 2%-isoflurane block ($p < .01$) and the effect was persistent after isoflurane termination during the 0%-isoflurane block ($p < .05$), whereas N6 peak time was significantly increased in not only the 2%- ($p < .01$) but also the 1.5%-isoflurane block ($p < .05$) with the effect persistent during the 0%-isoflurane block ($p < .05$) as well. Multiple comparison tests also revealed significant difference between the 2%- and the 0.5%-isoflurane blocks as well ($p < .05$). However, unlike auditory P3-to-P5 latency, visual P4-to-N6 latency was not affected by isoflurane.

Summary

Consistent with those previously reported in anesthetized human patients and animal models, we found a differentiating effect of isoflurane between AEPs and VEPs even when they were recorded simultaneously. In general, short-latency visually evoked potentials (SL-VEPs), short- and middle-latency auditory evoked potentials (SL-AEPs and ML-AEPs) appeared less susceptible to the suppressive effect of isoflurane (<20%) than long-latency auditory evoked potentials (LL-AEPs), middle- and long-latency visually evoked potentials (ML-VEPs and LL-

VEPs) (>30%). However, all the six AEPs or VEPs showed some changes during isoflurane anesthesia, which were statistically meaningful when choosing certain quantifications and components.

4.6. Discussions

Evoked potentials are averaged EEG signal time-locked to stimulus onsets. It is long-known that auditory evoked potentials (AEPs) can be viewed as multiple spectro-temporal components with different filtering parameters. These components, in spectral domain, were commonly divided into three categories and commonly termed as brainstem auditory evoked potentials (BAEPs), middle latency responses (MLRs), and cortical AEPs (cAEPs) (de la Torre, 2020; Maki, 2020; Shaw, 1988; Picton, 2010). In each of the three, multiple peak components can be further differentiated by their implicit time. For example, human BAEPs are composed of five positive peak components, named as wave I to wave V (Møller, 1998), and are generated by brainstem and midbrain auditory nuclei. Although cats and humans have nearly the same BAEPs in terms of their peak profiles (Buchwald & Huang, 1975; Fullerton et al., 1987; Lev & Sohmer, 1972), cat AEPs demonstrate more complexity beyond 10-ms (Dickerson & Buchwald, 1992; Farley & Starr, 1983; Starr & Farley, 1983) than human AEPs (Terence W Picton, 2010; Terence W Picton & Hillyard, 1974; Terry W Picton, Hillyard, Krausz, & Galambos, 1974). AEP studies using anesthetized animals, including the current study, in general only report the earliest cAEP components (P1 and N1) (Pincze, Lakatos, Rajkai, Ulbert, & Karmos, 2001, 2002; Presacco & Middlebrooks, 2018) but ignore the later components (e.g., P2 and N2) that are more identifiable in human ERP studies (Lightfoot, 2016; López-Caballero, Coffman, Seebold, Teichert, & Salisbury, 2022; Terence W Picton, Hillyard, Galambos, & Schiff, 1971; Ponton, Eggermont, Kwong, & Don, 2000; Woods, 1995). Similarly, visually evoked potentials are commonly studied

with components later than 75-ms, such as N1/N70, P1/P100, and N2/N140 (Creel, 2019; Di Russo et al., 2005; Sokol, 1976) whereas cat VEPs demonstrate response components with latencies as early as ~20-ms (Arakawa, Peachey, Celesia, & Rubboli, 1993). The pattern of cat VEPs is shown to differ depending on the stimulus (Aydin-Abidin, Moliadze, Eysel, & Funke, 2006; Sjöström, Abrahamsson, Norrsell, Helgason, & Roos, 1991) and age during early development (Ellingson & Wilcott, 1960).

The main purpose of our study was to examine the effect of isoflurane on AEPs/VEPs rather than the identification of the neural generator or making interpretations based on specific human AEP/VEP components. The labels used on individual peak components in the current study were arbitrary. Most of them followed a naming pattern of "polarity + implicit time". The exceptions are P3, P4, and P5, which corresponded to wave III, wave IV, and wave V from BAEP.

Effect of Isoflurane on Multi-modal, Spectrotemporal Components of Evoked Potentials

The present study is the first to document the evoked potentials of more than one sensory modality recorded simultaneously that were further processed for multiple spectro-temporal components (short-, middle-, and long-latency responses) under the same anesthesia conditions. Of all relevant previous studies, the closest one to the current study recorded BAEPs, cortical somatosensory evoked potentials (SSEPs) and VEPs from human patients in the same elective surgery (Sebel, Ingram, Flynn, Rutherfoord, & Rogers, 1986), where decreases in cortical SSEP and VEP amplitude were found but not BAER amplitudes. A similar study was carried out by the same group, with nitrous oxide used as anesthetic agent, and reported decrease in both SSEPs and VEPs amplitudes in healthy volunteers (Sebel, Flynn, & Ingram, 1984). In both studies, the evoked

potentials of three different modalities were measured sequentially at the end of 15-minute blocks of up-stepping isoflurane or N2O anesthesia (et-Iso: 0.55%, 1.1%, and 1.65%; inspired N2O: 10%, 30%, and 50%). The current study adopted a stimulus protocol that allows both AEPs and VEPs to be recorded within the same 60-second interval. By taking advantage of an ultrahigh sampling frequency (~6 kHz), we were also able to examine individual peak components, in addition to the overall waveforms.

In humans, a variety of anesthetics, such as isoflurane (Chi & Field, 1986; Chitranshi, You, Gupta, Klistorner, & Graham, 2015; Sloan, Sloan, & Rogers, 2010), enflurane (Chi & Field, 1990), sevoflurane (Ito et al., 2015; Tanaka et al., 2020; Tenenbein, Lam, Klein, & Lee, 2006), propofol (Tanaka et al., 2020; Tenenbein et al., 2006), thiopentone (Chi, Ryterband, & Field, 1989), and nitrous oxide (Sloan et al., 2010), have been shown to attenuate VEPs. Although it is unclear in cats whether different anesthetics affect scalp-recorded VEPs in the same way, one study examined VEPs derived from intracranial EEG under halothane, isoflurane, and enflurane anesthesia and identified that isoflurane, halothane, but not enflurane, suppressed VEPs recorded from visual cortex (Ogawa, Shingu, Shibata, Osawa, & Mori, 1992). Opposite to the consistent positive anesthetic effects identified in VEPs, it has been consistently reported that no anesthetic effects were found in SL-AEPs or BAEPs in humans with different agents, which include isoflurane (Manninen, Lam, & Nicholas, 1985; Sloan et al., 2010; Thornton et al., 1992), propofol (Nakagawa et al., 2006), thiopental (Drummond, Todd, & U, 1985), nitrous oxide (Sloan et al., 2010; Thornton et al., 1992), ketamine (Ruebhausen, Brozoski, & Bauer, 2012), but not sevoflurane (Nakagawa et al., 2006).

The effects of anesthetics on ML-AEP, on the other hand, has only been studied by a handful of groups. Two studies from the same group (Schwender, Kaiser, Klasing, Peter, &

Poppel, 1994; Schwender, Madler, Klasing, Peter, & Pöppel, 1994) reported that human ML-AEPs were completely absent under isoflurane, enflurane, thiopentone, etomidate, or propofol anesthesia, but unaffected under midazolam, diazepam, flunitrazepam, fentanyl, and ketamine anesthesia. The effect of isoflurane on ML-AEPs was graded by isoflurane concentration (Madler, Keller, Schwender, & Pöppel, 1991). Another group reported similar findings with enflurane (Thornton et al., 1983; Thornton, Heneghan, James, & Jones, 1984), halothane (Thornton et al., 1984), althesin (Thornton, Heneghan, Navaratnarajah, & Jones, 1986), etomidate (Thornton, Heneghan, Navaratnarajah, Bateman, & Jones, 1985), or isoflurane (Heneghan, Thornton, Navaratnarajah, & Jones, 1987), as well as propofol (Thornton et al., 1989). In the present study, however, even the highest concentrations of isoflurane did not achieve the absence of ML-AEPs as observed in humans. Instead, the increase of isoflurane concentration resulted in more waveform reformation rather than attenuation for ML-AEPs. This was shown by two previous reports in humans (Dutton, Smith, Rampil, Chortkoff, & Eger, 1999; Iselin-Chaves, El Moalem, Gan, Ginsberg, & Glass, 2000) using desflurane and propofol, and in dogs (Murrell, de Groot, Haagen, van den Brom, & Hellebrekers, 2004; Murrell, de Groot, Psatha, & Hellebrekers, 2005) using sevoflurane and acepromazine.

LL-AEPs also demonstrated more waveform alteration than attenuation, although our data showed significant decrease in the signal strength of LL-AEPs during isoflurane blocks. It was also found in rats that P1-to-N1 amplitudes were decreased by 1.5 - 2% isoflurane (Brewer et al., 2021). However, it was known that human LL-AEP components N1 and P2 were almost abolished after anesthesia induction with various anesthetic agents (Plourde & Boylan, 1991; Plourde & Picton, 1991; Simpson et al., 2002). Although it is unclear why early components of animal LL-

AEPs are more resilient to anesthesia than human LL-AEPs, these response components may serve as useful models for studying the effect of anesthetics on sensory processing.

Finally, the short latency components found in VEPs, which were time-locked to flash onset and referred as short-latency VEPs or SL-VEPs, were indeed decreased by isoflurane and increased after isoflurane administration terminated. Since isoflurane is administrated through an electrically-insulated system, its effect on SL-VEPs, consistent across subjects, can only be physiological and thereby indicating SL-VEPs themselves are unlikely to be an electrical artifact. Most studies suggested that the earliest visually evoked response has a latency of more than 20-ms, although one study reported some response latencies as short as 12 ms (Cracco & Cracco, 1978). In addition, the waveform of our SL-VEPs resembled the spike waveform recorded from optic nerve using electrical stimulation, which has a latency as brief as 1- to 3-msec (Bishop, 1953; Creutzfeldt & Kuhnt, 1973). We identified the earliest visual peak component at ~4 ms after flash onset. The waveform was different from those SL-AEPs we recorded, which could exclude the possibility of the contamination by an auditory response.

Dose-dependent Effect of Isoflurane on Sensory Processing

A secondary factor that we investigated was the relationship of isoflurane dosage and its effect on auditory and/or visually evoked potentials (**Fig. 6**). For most measurements derived from the AEP/VEP waveforms that are with a significant main block effect, the multiple comparison tests afterwards revealed one to three pairs of comparisons with a significant difference usually involving 2%-isoflurane block. This suggested that isoflurane likely needs to reach a concentration of 2% to make a statistically meaningful change to AEPs and VEPs. However, isoflurane did achieve significant change at a concentration of 1.5% in some of those measurements (**Fig. 6 BC,**

underscored), including auditory P20 peak time, auditory P18 amplitude and peak time, auditory N21 peak amplitude, auditory P3 and P5 peak time, and visual N6 peak time. Although it seemed that most 1.5%-isoflurane effect was found in AEPs with the only exception of VEP component being visual N6 peak time, we cannot draw the conclusion that AEPs was susceptible to isoflurane of lower concentration than VEPs. These observations, however, could have shed some light on the uncharted dose-dependency of isoflurane effect on AEPs and VEPs that have yet to be fully investigated.

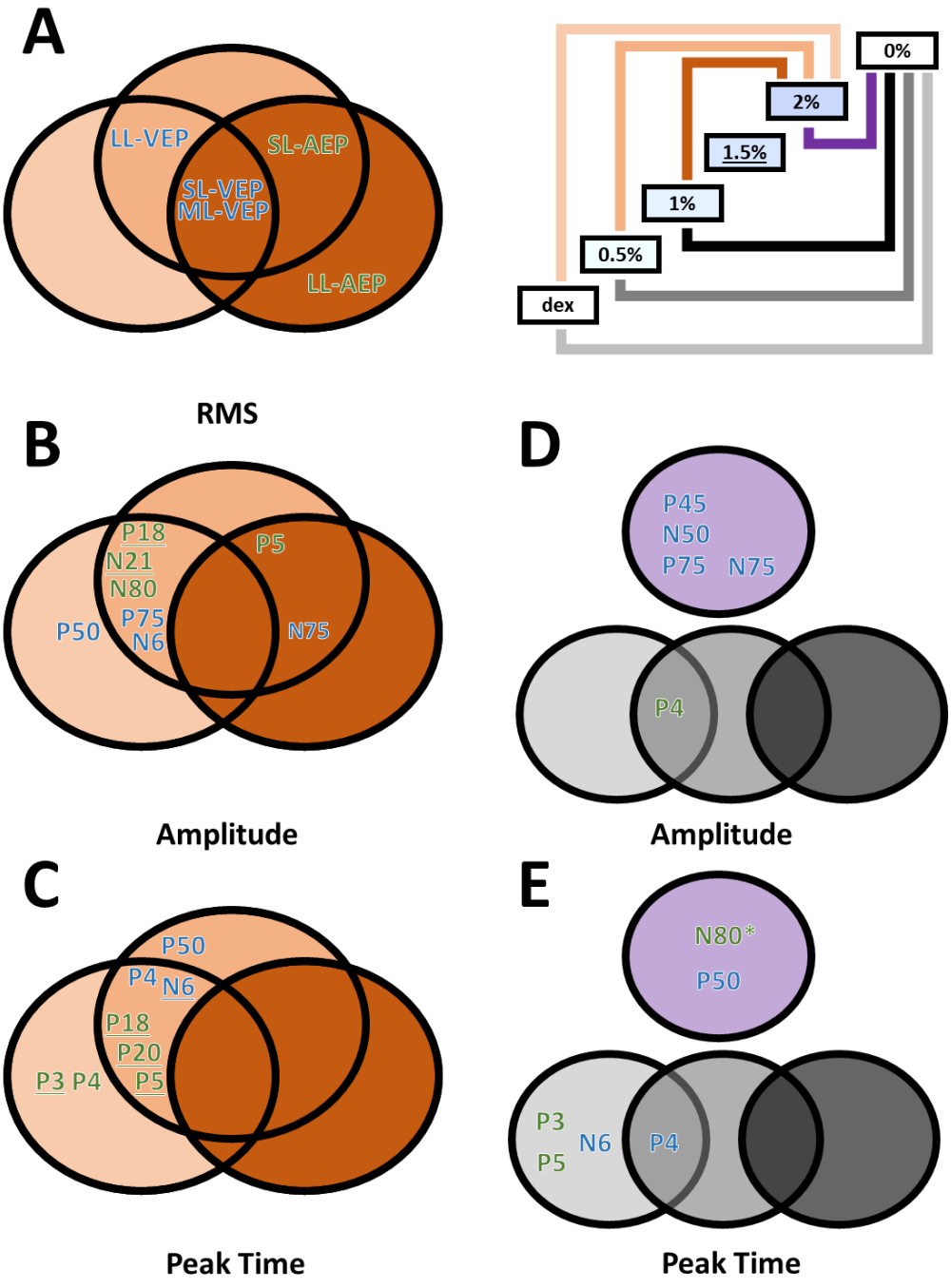


Figure 6. Venn diagrams of dose-dependent and wash-out effect of isoflurane in different measurements. (A-C) Effect of isoflurane when compared to the 2%-isoflurane blocks. Green, AEP measurements. Blue, VEP measurements. Underlined, additionally significant difference between the 1.5%-isoflurane and the dexmedetomidine blocks. (D-E) Wash-out effect of isoflurane during the 0%-isoflurane block. * N80 peak time also showed significant difference between the 0%- and the 1.5%-isoflurane blocks.

The washout effect after the termination of isoflurane was observed. For VEPs, the comparison between the 2%-block and the following 0%-isoflurane block revealed that the amplitudes of all three LL-VEP components P45, N50, P75, as well as the P50 peak time and the N75 peak amplitude from ML-VEP, significantly recovered in the first 10 minutes after the subjects stopped receiving isoflurane. The only AEP component that demonstrated the same washout effect was N80 peak time from LL-AEPs. The difference between AEPs and VEPs in the washout effect suggested that, without further isoflurane administration, the level of isoflurane in visual cortical areas was not sustained for as long as compared to auditory areas. This may explain an earlier report of intraoperative VEP monitoring data where VEPs were not completely absent, but very unreliable (Kamio et al., 2014). In the current study, we also observed the opposite of the washout effect in some components from SL-AEP and SL-VEP, where the difference between 0%-isoflurane block and one of the low-concentration (dexmedetomidine, i.e., Dex, 0.5%- or 1%-isoflurane) blocks is statistically significant (**Fig. 6 DE**). The lack of recovery may indicate that the effect observed during high-concentration isoflurane blocks may be due to non-isoflurane factors, such as body temperature or the wear-off of dexmedetomidine.

There was a noticeable trend in our data that there was less variance across subjects in the 0.5%-isoflurane block than the Dex block. The smaller data variance led to a larger statistical power in 0.5%-isoflurane block when it was compared to 2%-isoflurane block. In particular, multiple comparison tests did not detect a significant difference between the 2%-isoflurane and the Dex block, but between the 2%- and the 0.5%-isoflurane block, for the RMS values of SL-AEPs, the P50 peak time and the N75 amplitude from ML-VEPs, and the P5 amplitude from SL-AEPs (**Fig. 6 ABC**). Although dexmedetomidine is known to not affect evoked potentials (Rozet et al., 2015), the action it has on the locus coeruleus, a central adrenergic nucleus, may trigger rat

epileptic activities, which can be suppressed by co-administrating isoflurane at a low concentration (~0.3%) (Fukuda, Vazquez, Zong, & Kim, 2013). Compared to midazolam sedation, human auditory evoked potentials appeared more variable (Haenggi et al., 2006). Therefore, our observation on data variance may reflect the synergistic effect of dexmedetomidine and low-concentration isoflurane. This should be considered by future studies of both drugs, especially for the selection of baseline, given the fact that dexmedetomidine is frequently administrated prior to anesthesia in the clinical setting (Kaur & Singh, 2011; Naaz & Ozair, 2014; Sun, Lu, Huang, & Jiang, 2014).

In short, the dosage manipulation of the four inspired concentrations, i.e., 0.5%, 1%, 1.5%, and 2%, in our study revealed statistically significant effects of isoflurane on AEPs and VEPs, contrasting the 1.5%- or the 2%-isoflurane block with the Dex or the 0.5%-isoflurane block. Cats pre-medicated with dexmedetomidine with a sample size as small as five were enough to demonstrate the dose-dependent effect of isoflurane on AEPs and VEPs, as well as its wash-out effect.

Mechanism of Isoflurane

Although surveys showed that sevoflurane has been becoming more commonly preferred by anesthetists in recent years, isoflurane was widely used in clinical practice for >30 years since it was first approved by FDA in 1979. For more than 30 years, a vast amount of research on isoflurane has been carried out in human, in vivo animal models, and in vitro cultured cells. The mechanism of isoflurane as an anesthetic, and as a confounding factor in neurophysiology experiments using anesthetized animals as well, is still being investigated today.

$GABA_A$ receptors are deemed as the primary target of isoflurane, like many other intravenous and volatile anesthetics. At low concentrations, isoflurane is known to potentiate GABA-induced chloride current (Harrison, Kugler, Jones, Greenblatt, & Pritchett, 1993; Krasowski et al., 1998; Nakahiro, Yeh, Brunner, & Narahashi, 1989; Zhang, Stabernack, Sonner, Dutton, & Eger, 2001). Like other anesthetics, it was shown that isoflurane, under specific conditions, can directly induce hyperpolarizing chloride current through GABA receptors directly (Jia et al., 2008; Yang, Isenberg, & Zorumski, 1992). Either direct or indirect facilitation of GABAergic inhibition caused by isoflurane may serve as the major underlying mechanism of its effect on AEPs and VEPs.

The effect of GABAergic inhibition on auditory evoked potentials has been shown along rat ascending auditory pathway by systematically administrated barbiturates, which are known to have an agnostic effect on GABAergic receptors. Evoked potentials recorded from the cochlear nucleus and the inferior colliculus were substantially suppressed during pentobarbital anesthesia (40 mg/kg) when compared to awake recording (Webster & Aitkin, 1971). In auditory thalamus, however, it was shown that pentobarbital (35 – 40 mg/kg) causes less suppression when compared to auditory cortex (Mäkelä, Karmos, Molnar, Csepe, & Winkler, 1990), or even facilitation (Webster & Aitkin, 1971). In rat auditory thalamus or primary auditory cortex, the effect of bolus pentobarbital on neuronal activities were studied in rats under ketamine (Wehr & Zador, 2005) or nitrous oxide anesthesia (Zurita, Villa, De Ribaupierre, De Ribaupierre, & Rouiller, 1994). Intracellular recording demonstrated a dramatic prolongation of sound-induced inhibitory conductance caused by bolus pentobarbital (50 mg/kg) (Wehr & Zador, 2005), which suggests a large involvement of GABAergic inhibition in the auditory response generated in cortex.

The effect of GABAergic agents on visual responses were also shown in retinal rod bipolar cells (Eggers & Lukasiewicz, 2006), retinal ganglion neurons (Priest, Robbins, & Ikeda, 1985), and visual cortex (Kraut, Arezzo, & Vaughan Jr, 1990). Although our results showed that ML- and LL-VEPs were substantially attenuated during the 2%-isoflurane block, we did not achieve as much suppression in AEPs using isoflurane as expected from the barbiturate studies. Aside from the likely deeper anesthesia using pentobarbital, the less suppression in AEPs in the current study is also possible as isoflurane may work through a non-GABAergic mechanism.

It was shown that isoflurane acts through a different glutamatergic mechanism when compared to sevoflurane. Although both isoflurane and sevoflurane reduced presynaptic calcium-dependent glutamate release, it was only isoflurane but not sevoflurane that showed a decrease in the reduction as its dose increased (Larsen, Valø, Berg-Johnsen, & Langmoen, 1998; Vinje, Moe, Valø, & Berg-Johnsen, 2002). EEG recording under isoflurane but not sevoflurane anesthesia showed a decreased beta-band energy, and it was proposed that the glutamatergic mechanism of isoflurane is more complicated than that of sevoflurane and involves more than one signal pathway (Arakawa et al., 1993). It is still unclear whether this glutamatergic mechanism of isoflurane is also involved in its effect on AEPs and VEPs.

Volatile anesthetics including isoflurane are also known to have additional targets, such as two-pore-domain K+ channels (Patel et al., 1999) and NMDA receptors (Dickinson et al., 2007; Stabernack et al., 2003). The interpretation of electrophysiology recording made during isoflurane anesthesia, either extracellular recording (Noda & Takahashi, 2015; White, Abbott, & Fiser, 2012) or evoked potentials like the current study, were complicated by multi-target mechanism of isoflurane.

Comparison between auditory and visual system

As far as we know, there is very little discussion on how the effect of anesthetics is different across different sensory modalities. The auditory and the visual systems have been compared from various perspectives, including cortical columnar distribution and thalamocortical transformation (Linden & Schreiner, 2003), parallel what- and where-pathways (Rauschecker, 2015), and auditory and visual scene analysis (Kondo, van Loon, Kawahara, & Moore, 2017). Along ascending auditory pathway, acoustic input initially coded by the cochlea is processed in multiple stations, including the brainstem, midbrain, thalamic, and cortical nucleus/regions (Malmierca & Hackett, 2010).

In comparison, visual inputs are initially coded and processed at retina, then carried to the thalamus by retinal ganglion neurons, and eventually relayed to visual cortex (Dacheux & Miller, 1981; Masland, 2012; Wurtz & Kandel, 2000). If the effect of anesthetics were applied globally rather than specifically to certain brain regions, one would expect to see more drastic changes in AEPs, especially late cortical components, than in VEPs, as the auditory ascending pathway requires more synaptic relays. This was not the case in our results. In fact, we observed more resistance to isoflurane in AEPs than VEPs, suggesting that synapses along ascending auditory and visual pathways are not equivalently affected by isoflurane, at least for light anesthesia.

The susceptibility of the visual system was not only observed with electrophysiology during intraoperative monitoring and anesthesia trial, but also with brain imaging techniques. There are already several MRI or PET studies on the effect of anesthetic agents on auditory and visual functions (Hudetz, 2012). Using resting-state functional MRI, Boveroux et al. (2010)

compared the networks seeded by primary auditory and visual cortices between awake and unconscious subjects, and found no effect of propofol on functional connectivity within either one of the networks but a significant disruption in the correlation between the two. With the advancement of brain imaging techniques in animal models (Rohlfing et al., 2012; Stolzberg, Wong, Butler, & Lomber, 2017), the differences between the auditory and the visual systems suggested by our findings may be further investigated.

Summary and implications for future studies

In this study, we examined the effect of isoflurane on auditory evoked potentials (AEPs) and visually evoked potentials (VEPs). The effect of isoflurane on signal strength prominently differentiated LL- and ML-VEPs, whereas its effect on individual peak components revealed a variety of patterns that were complicated by dose-dependency. Such variation may be explained by the complex cellular and molecular mechanism of isoflurane and therefore calls for more future research incorporating both electrophysiology and brain imaging techniques. The findings reported here can benefit from future replications using different stimulus rates for clicks and flashes as well as other stimulation devices and recording media. It is known that responses in different regions along the ascending auditory and visual pathways display different preferences of stimulus rate (Eggermont, 1991; Jones, 1980; Langner & Schreiner, 1988). LEDs-attached goggles have been used in surgery rooms for monitoring VEPs in patients (Kodama et al., 2010; Soffin et al., 2018). There are other options of recording electrodes aside from subdermal needles used in the current study, such as bone screws and cup electrodes, which may produce AEPs and VEPs with less noise (Santangelo et al., 2018; You, Klistorner, Thie, & Graham, 2011) and additional spatial information about their neural generators (Plourde et al., 2008). On the other

hand, the manipulation of isoflurane concentration can be improved by titrating for end-tidal concentration instead of controlling inspired concentration.

The use of sensory evoked potentials in intraoperative monitoring is promising but requires improvements in protocols and devices. Currently, there are three EEG-based quantifications, including bispectral index (BIS), E-Entrophy, and Narcotrend-Compact M, recommended in clinical settings for monitoring anesthesia depth (National Institute for Health and Care Excellence, 2012). The use of AEPs and VEPs may serve as an augmentation in EEG-based anesthesia monitoring, as it allows for a continuous examination on the functional integrity along the ascending auditory and visual pathways in patients undergoing neurosurgery.

4.7. Reference

Arakawa, K., Peachey, N. S., Celesia, G. G., & Rubboli, G. (1993). Component-specific effects of physostigmine on the cat visual evoked potential. *Experimental brain research, 95*(2), 271-276.

Aydin-Abidin, S., Moliadze, V., Eysel, U. T., & Funke, K. (2006). Effects of repetitive TMS on visually evoked potentials and EEG in the anaesthetized cat: dependence on stimulus frequency and train duration. *The Journal of physiology, 574*(2), 443-455.

Banoub, M., Tetzlaff, J. E., & Schubert, A. (2003). Pharmacologic and physiologic influences affecting sensory evoked potentials: implications for perioperative monitoring. *The Journal of the American Society of Anesthesiologists, 99*(3), 716-737.

Bishop, P. O. (1953). Synaptic transmission. An analysis of the electrical activity of the lateral geniculate nucleus in the cat after optic nerve stimulation. *Proceedings of the Royal Society of London. Series B-Biological Sciences, 141*(904), 362-392.

Boveroux, P., Vanhaudenhuyse, A., Bruno, M.-A., Noirhomme, Q., Lauwick, S., Luxen, A., . . . Phillips, C. (2010). Breakdown of within-and between-network resting state functional magnetic resonance imaging connectivity during propofol-induced loss of consciousness. *The Journal of the American Society of Anesthesiologists, 113*(5), 1038-1053.

Brewer, L., Holdford, M., Holloway, Z., Sable, J., Andrasik, F., & Sable, H. (2021). Isoflurane effects on the N1 and other long-latency auditory evoked potentials in Wistar rats. *Neuroscience Research, 173*, 71-79.

Buchwald, J. S., & Huang, C.-M. (1975). Far-field acoustic response: origins in the cat. *Science, 189*(4200), 382-384.

Caron-Desrochers, L., Schönwiesner, M., Focke, K., & Lehmann, A. (2018). Assessing visual modulation along the human subcortical auditory pathway. *Neuroscience letters, 685*, 12-17.

Chi, O. Z., & Field, C. (1986). Effects of isoflurane on visual evoked potentials in humans. *Anesthesiology (Philadelphia), 65*(3), 328-330.

Chi, O. Z., & Field, C. (1990). Effects of enflurane on visual evoked potentials in humans. *British Journal of Anaesthesia, 64*(2), 163-166.

Chi, O. Z., Ryterband, S., & Field, C. (1989). Visual evoked potentials during thiopentone-fentanylnitrous oxide anaesthesia in humans. *Canadian journal of anaesthesia, 36*(6), 637-640.

Chitranshi, N., You, Y., Gupta, V. K., Klistorner, A., & Graham, S. L. (2015). Effects of isoflurane on the visual evoked potentials in rats. *Investigative Ophthalmology & Visual Science, 56*(7), 468-468.

Cracco, R. Q., & Cracco, J. B. (1978). Visual evoked potential in man: early oscillatory potentials. *Electroencephalography and clinical neurophysiology, 45*(6), 731-739.

Creel, D. J. (2019). Visually evoked potentials. *Handbook of clinical neurology, 160*, 501-522.

Creutzfeldt, O. D., & Kuhnt, U. (1973). Electrophysiology and topographical distribution of visual evoked potentials in animals. In *Visual Centers in the Brain* (pp. 595-646): Springer.

Dacheux, R. F., & Miller, R. F. (1981). An intracellular electrophysiological study of the ontogeny of functional synapses in the rabbit retina. II. Amacrine cells. *Journal of Comparative Neurology, 198*(2), 327-334.

Di Russo, F., Pitzalis, S., Spitoni, G., Aprile, T., Patria, F., Spinelli, D., & Hillyard, S. A. (2005). Identification of the neural sources of the pattern-reversal VEP. *Neuroimage, 24*(3), 874-886.

Dickerson, L. W., & Buchwald, J. S. (1992). Long-latency auditory-evoked potentials: role of polysensory association cortex in the cat. *Experimental neurology, 117*(3), 313-324.

Dickinson, R., Peterson, B. K., Banks, P., Simillis, C., Martin, J. C. S., Valenzuela, C. A., . . . Franks, N. P. (2007). Competitive inhibition at the glycine site of the N-methyl-D-aspartate receptor by the anesthetics xenon and isoflurane: evidence from molecular modeling and electrophysiology. *The Journal of the American Society of Anesthesiologists, 107*(5), 756-767.

Drummond, J. C., Todd, M. M., & U, H. S. (1985). The effect of high dose sodium thiopental on brain stem auditory and median nerve somatosensory evoked responses in humans. *The Journal of the American Society of Anesthesiologists, 63*(3), 249-254.

Dutton, R. C., Smith, W. D., Rampil, I. J., Chortkoff, B. S., & Eger, E. I. (1999). Forty-hertz midlatency auditory evoked potential activity predicts wakeful response during desflurane and propofol anesthesia in volunteers. *The Journal of the American Society of Anesthesiologists, 91*(5), 1209-1209.

Eggermont, J. J. (1991). Rate and synchronization measures of periodicity coding in cat primary auditory cortex. *Hearing Research, 56*(1-2), 153-167.

Eggers, E. D., & Lukasiewicz, P. D. (2006). GABAA, GABAC and glycine receptor-mediated inhibition differentially affects light-evoked signalling from mouse retinal rod bipolar cells. *The Journal of physiology, 572*(1), 215-225.

Ellingson, R. J., & Wilcott, R. C. (1960). Development of evoked responses in visual and auditory cortices of kittens. *Journal of neurophysiology, 23*(4), 363-375.

Excellence, N. I. f. H. a. C. (2012). Depth of anaesthesia monitors-Bispectral index (BIS), E--Entropy and Narcotrend Compact M

Farley, G. R., & Starr, A. (1983). Middle and long latency auditory evoked potentials in cat. I. Component definition and dependence on behavioral factors. *Hearing Research, 10*(2), 117-138.

Fukuda, M., Vazquez, A. L., Zong, X., & Kim, S. G. (2013). Effects of the α2-adrenergic receptor agonist dexmedetomidine on neural, vascular and BOLD fMRI responses in the somatosensory cortex. *European Journal of Neuroscience, 37*(1), 80-95.

Fullerton, B. C., Levine, R. A., Hosford-Dunn, H. L., & Kiang, N. Y. (1987). Comparison of cat and human brain-stem auditory evoked potentials. *Electroencephalography and clinical neurophysiology, 66*(6), 547-570.

Grasshoff, C., & Antkowiak, B. (2006). Effects of isoflurane and enflurane on GABAA and glycine receptors contribute equally to depressant actions on spinal ventral horn neurones in rats. *BJA: British Journal of Anaesthesia, 97*(5), 687-694.

Haenggi, M., Ypparila, H., Hauser, K., Caviezel, C., Korhonen, I., Takala, J., & Jakob, S. M. (2006). The effects of dexmedetomidine/remifentanil and midazolam/remifentanil on auditory-evoked potentials and electroencephalogram at light-to-moderate sedation levels in healthy subjects. *Anesthesia & Analgesia, 103*(5), 1163-1169.

Harrison, N. L., Kugler, J., Jones, M., Greenblatt, E., & Pritchett, D. (1993). Positive modulation of human gamma-aminobutyric acid type A and glycine receptors by the inhalation anesthetic isoflurane. *Molecular pharmacology, 44*(3), 628-632.

Heneghan, C., Thornton, C., Navaratnarajah, M., & Jones, J. (1987). Effect of isoflurane on the auditory evoked response in man. *BJA: British Journal of Anaesthesia, 59*(3), 277-282.

Hudetz, A. G. (2012). General anesthesia and human brain connectivity. *Brain connectivity, 2*(6), 291-302.

Iselin-Chaves, I. A., El Moalem, H. E., Gan, T. J., Ginsberg, B., & Glass, P. S. (2000). Changes in the auditory evoked potentials and the bispectral index following propofol or propofol and alfentanil. *The Journal of the American Society of Anesthesiologists, 92*(5), 1300-1310.

Ito, Y., Maehara, S., Itoh, Y., Hayashi, M., Kubo, A., Itami, T., . . . Yamashita, K. (2015). Effect of sevoflurane concentration on visual evoked potentials with pattern stimulation in dogs. *Journal of Veterinary Medical Science, 77*(2), 155-160.

Jehle, T., Ehlken, D., Wingert, K., Feuerstein, T., Bach, M., & Lagreze, W. (2009). Influence of narcotics on luminance and frequency modulated visual evoked potentials in rats. *Documenta ophthalmologica, 118*(3), 217-224.

Jia, F., Yue, M., Chandra, D., Homanics, G. E., Goldstein, P. A., & Harrison, N. L. (2008). Isoflurane is a potent modulator of extrasynaptic GABAA receptors in the thalamus. *Journal of Pharmacology and Experimental Therapeutics, 324*(3), 1127-1135.

Jones, K. R. (1980). *Effects of Visual Deprivation on the Temporal Modulation Sensitivity of the Optic Tract, Lateral Geniculate Nucleus, and Visual Cortex of the Cat*: The Florida State University.

Kamio, Y., Sakai, N., Sameshima, T., Takahashi, G., Koizumi, S., Sugiyama, K., & Namba, H. (2014). Usefulness of intraoperative monitoring of visual evoked potentials in transsphenoidal surgery. *Neurologia medico-chirurgica*, oa. 2014-0023.

Kaur, M., & Singh, P. (2011). Current role of dexmedetomidine in clinical anesthesia and intensive care. *Anesthesia, essays and researches*, *5*(2), 128.

Kodama, K., Goto, T., Sato, A., Sakai, K., Tanaka, Y., & Hongo, K. (2010). Standard and limitation of intraoperative monitoring of the visual evoked potential. *Acta neurochirurgica, 152*(4), 643-648.

Kondo, H. M., van Loon, A. M., Kawahara, J.-I., & Moore, B. C. (2017). Auditory and visual scene analysis: an overview. *Philosophical Transactions of the Royal Society B: Biological Sciences, 372*(1714), 20160099.

Krasowski, M. D., Koltchine, V. V., Rick, C. E., Ye, Q., Finn, S. E., & Harrison, N. L. (1998). Propofol and other intravenous anesthetics have sites of action on the γ-aminobutyric acid type A receptor distinct from that for isoflurane. *Molecular pharmacology, 53*(3), 530-538.

Kraut, M. A., Arezzo, J. C., & Vaughan Jr, H. G. (1990). Inhibitory processes in the flash evoked potential of the monkey. *Electroencephalography and clinical neurophysiology, 76*(5), 440-452.

Langner, G., & Schreiner, C. E. (1988). Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. *Journal of neurophysiology, 60*(6), 1799-1822.

Larsen, M., Valø, E., Berg-Johnsen, J., & Langmoen, I. (1998). Isoflurane reduces synaptic glutamate release without changing cytosolic free calcium in isolated nerve terminals. *European journal of anaesthesiology, 15*(2), 224-229.

Lev, A., & Sohmer, H. (1972). Sources of averaged neural responses recorded in animal and human subjects during cochlear audiometry (electro-cochleogram). *Archiv für klinische und experimentelle Ohren-, Nasen-und Kehlkopfheilkunde, 201*(2), 79-90.

Lightfoot, G. (2016). *Summary of the N1-P2 cortical auditory evoked potential to estimate the auditory threshold in adults.* Paper presented at the Seminars in hearing.

Linden, J. F., & Schreiner, C. E. (2003). Columnar transformations in auditory cortex? A comparison to visual and somatosensory cortices. *Cerebral cortex, 13*(1), 83-89.

López-Caballero, F., Coffman, B., Seebold, D., Teichert, T., & Salisbury, D. F. (2022). Intensity and inter-stimulus-interval effects on human middle-and long-latency auditory evoked potentials in an unpredictable auditory context. *Psychophysiology*, e14217.

Madler, C., Keller, I., Schwender, D., & Pöppel, E. (1991). Sensory information processing during general anaesthesia: effect of isoflurane on auditory evoked neuronal oscillations. *British Journal of Anaesthesia, 66*(1), 81-87.

Mäkelä, J., Karmos, G., Molnar, M., Csepe, V., & Winkler, I. (1990). Steady-state responses from the cat auditory cortex. *Hearing Research, 45*(1-2), 41-50.

Malmierca, M. S., & Hackett, T. A. (2010). Structural organization of the ascending auditory pathway. *The Auditory Brain, 1147*, 9-41.

Manninen, P. H., Lam, A. M., & Nicholas, J. F. (1985). The effects of isoflurane and isoflurane-nitrous oxide anesthesia on brainstem auditory evoked potentials in humans. *Anesthesia and analgesia, 64*(1), 43-47.

Masland, R. H. (2012). The tasks of amacrine cells. *Visual neuroscience, 29*(1), 3-9.

Møller, A. R. (1998). *Neural generators of the brainstem auditory evoked potentials.* Paper presented at the Seminars in hearing.

Murrell, J. C., de Groot, H., Haagen, A. V. v., van den Brom, W., & Hellebrekers, L. (2004). Middle-latency auditory-evoked potential in acepromazine-sedated dogs. *Journal of veterinary internal medicine, 18*(2), 196-200.

Murrell, J. C., de Groot, H. N., Psatha, E., & Hellebrekers, L. J. (2005). Investigation of changes in the middle latency auditory evoked potential during anesthesia with sevoflurane in dogs. *American journal of veterinary research, 66*(7), 1156-1161.

Naaz, S., & Ozair, E. (2014). Dexmedetomidine in current anaesthesia practice-a review. *Journal of clinical and diagnostic research: JCDR, 8*(10), GE01.

Nakagawa, I., Hidaka, S., Okada, H., Kubo, T., Okamura, K., & Kato, T. (2006). Effects of sevoflurane and propofol on evoked potentials during neurosurgical anesthesia. *Masui. The Japanese Journal of Anesthesiology, 55*(6), 692-698.

Nakahiro, M., Yeh, J. Z., Brunner, E., & Narahashi, T. (1989). General anesthetics modulate GABA receptor channel complex in rat dorsal root ganglion neurons. *The FASEB journal, 3*(7), 1850-1854.

Noda, T., & Takahashi, H. (2015). Anesthetic effects of isoflurane on the tonotopic map and neuronal population activity in the rat auditory cortex. *European Journal of Neuroscience, 42*(6), 2298-2311.

Ogawa, T., Shingu, K., Shibata, M., Osawa, M., & Mori, K. (1992). The divergent actions of volatile anaesthetics on background neuronal activity and reactive capability in the central nervous system in cats. *Canadian journal of anaesthesia, 39*(8), 862-872.

Patel, A. J., Honoré, E., Lesage, F., Fink, M., Romey, G., & Lazdunski, M. (1999). Inhalational anesthetics activate two-pore-domain background K+ channels. *Nature neuroscience, 2*(5), 422-426.

Perry, V., & Cowey, A. (1984). Retinal ganglion cells that project to the superior colliculus and pretectum in the macaque monkey. *Neuroscience, 12*(4), 1125-1137.

Picton, T. W. (2010). *Human auditory evoked potentials*: Plural Publishing.

Picton, T. W., & Hillyard, S. A. (1974). Human auditory evoked potentials. II: Effects of attention. *Electroencephalography and clinical neurophysiology, 36*, 191-200.

Picton, T. W., Hillyard, S. A., Galambos, R., & Schiff, M. (1971). Human auditory attention: a central or peripheral process? *Science, 173*(3994), 351-353.

Picton, T. W., Hillyard, S. A., Krausz, H. I., & Galambos, R. (1974). Human auditory evoked potentials. I: Evaluation of components. *Electroencephalography and clinical neurophysiology, 36*, 179-190.

Pincze, Z., Lakatos, P., Rajkai, C., Ulbert, I., & Karmos, G. (2001). Separation of mismatch negativity and the N1 wave in the auditory cortex of the cat: a topographic study. *Clinical neurophysiology, 112*(5), 778-784.

Pincze, Z., Lakatos, P., Rajkai, C., Ulbert, I., & Karmos, G. (2002). Effect of deviant probability and interstimulus/interdeviant interval on the auditory N1 and mismatch negativity in the cat auditory cortex. *Cognitive Brain Research, 13*(2), 249-253.

Plourde, G., & Boylan, J. F. (1991). The long-latency auditory evoked potential as a measure of the level of consciousness during sufentanil anesthesia. *Journal of cardiothoracic and vascular anesthesia, 5*(6), 577-583.

Plourde, G., Garcia-Asensi, A., Backman, S., Deschamps, A., Chartrand, D., Fiset, P., & Picton, T. W. (2008). Attenuation of the 40-hertz auditory steady state response by propofol involves the cortical and subcortical generators. *The Journal of the American Society of Anesthesiologists, 108*(2), 233-242.

Plourde, G., & Picton, T. (1991). Long-latency auditory evoked potentials during general anesthesia: N1 and P3 components. *Anesthesia and analgesia, 72*(3), 342-350.

Ponton, C. W., Eggermont, J. J., Kwong, B., & Don, M. (2000). Maturation of human central auditory system activity: evidence from multi-channel evoked potentials. *Clinical neurophysiology, 111*(2), 220-236.

Presacco, A., & Middlebrooks, J. C. (2018). Tone-evoked acoustic change complex (ACC) recorded in a sedated animal model. *Journal of the Association for Research in Otolaryngology, 19*(4), 451-466.

Priest, T., Robbins, J., & Ikeda, H. (1985). The action of inhibitory neurotransmitters, γ-aminobutyric acid and glycine may distinguish between the area centralis and the peripheral retina in cats. *Vision research, 25*(12), 1761-1770.

Rauschecker, J. P. (2015). Auditory and visual cortex of primates: a comparison of two sensory systems. *European Journal of Neuroscience, 41*(5), 579-585.

Rohlfing, T., Kroenke, C. D., Sullivan, E. V., Dubach, M. F., Bowden, D. M., Grant, K. A., & Pfefferbaum, A. (2012). The INIA19 template and NeuroMaps atlas for primate brain image parcellation and spatial normalization. *Frontiers in neuroinformatics, 6*, 27.

Rozet, I., Metzner, J., Brown, M., Treggiari, M. M., Slimp, J. C., Kinney, G., . . . Vavilala, M. S. (2015). Dexmedetomidine does not affect evoked potentials during spine surgery. *Anesthesia & Analgesia, 121*(2), 492-501.

Ruebhausen, M., Brozoski, T., & Bauer, C. (2012). A comparison of the effects of isoflurane and ketamine anesthesia on auditory brainstem response (ABR) thresholds in rats. *Hearing Research, 287*(1-2), 25-29.

Samra, S. K., Vanderzant, C. W., Domer, P. A., & Sackellares, J. C. (1987). Differential effects of isoflurane on human median nerve somatosensory evoked potentials. *The Journal of the American Society of Anesthesiologists, 66*(1), 29-35.

Santangelo, R., Castoldi, V., D'Isa, R., Marenna, S., Huang, S.-C., Cursi, M., . . . Leocani, L. (2018). Visual evoked potentials can be reliably recorded using noninvasive epidermal electrodes in the anesthetized rat. *Documenta ophthalmologica, 136*(3), 165-175.

Schwender, D., Kaiser, A., Klasing, S., Peter, K., & Poppel, E. (1994). Midlatency auditory evoked potentials and explicit and implicit memory in patients undergoing cardiac surgery. *Anesthesiology, 80*(3), 493-501. doi:10.1097/00000542-199403000-00004

Schwender, D., Madler, C., Klasing, S., Peter, K., & Pöppel, E. (1994). Anesthetic control of 40-Hz brain activity and implicit memory. *Consciousness and Cognition, 3*(2), 129-147.

Sebel, P., Erwin, C., & Neville, W. (1987). Effects of halothane and enflurane on far and near field somatosensory evoked potentials. *BJA: British Journal of Anaesthesia, 59*(12), 1492-1496.

Sebel, P., Flynn, P., & Ingram, D. (1984). Effect of nitrous oxide on visual, auditory and somatosensory evoked potentials. *British Journal of Anaesthesia, 56*(12), 1403-1407.

Sebel, P., Ingram, D., Flynn, P., Rutherfoord, C., & Rogers, H. (1986). Evoked potentials during isoflurane anaesthesia. *BJA: British Journal of Anaesthesia, 58*(6), 580-585.

Simpson, T., Manara, A., Kane, N., Barton, R., Rowlands, C., & Butler, S. (2002). Effect of propofol anaesthesia on the event-related potential mismatch negativity and the auditory-evoked potential N1. *British Journal of Anaesthesia, 89*(3), 382-388.

Sjöström, A., Abrahamsson, M., Norrsell, K., Helgason, G., & Roos, A. (1991). Flashed pattern-induced activity in the visual system: I. The short latency evoked response recorded from the cat visual cortex. *Acta physiologica scandinavica, 143*(1), 1-9.

Sloan, T., & Heyer, E. (2002). Anesthesia for intraoperative neurophysiologic monitoring of the spinal cord. *Journal of Clinical Neurophysiology, 19*(5), 430-443.

Sloan, T., Sloan, H., & Rogers, J. (2010). Nitrous oxide and isoflurane are synergistic with respect to amplitude and latency effects on sensory evoked potentials. *Journal of clinical monitoring and computing, 24*(2), 113-123.

Soffin, E., Emerson, R., Cheng, J., Mercado, K., Smith, K., & Beckman, J. (2018). A pilot study to record visual evoked potentials during prone spine surgery using the SightSaver™ photic visual stimulator. *Journal of clinical monitoring and computing, 32*(5), 889-895.

Sokol, S. (1976). Visually evoked potentials: theory, techniques and clinical applications. *Survey of ophthalmology, 21*(1), 18-44.

Stabernack, C., Sonner, J. M., Laster, M., Zhang, Y., Xing, Y., Sharma, M., & Eger, E. I. (2003). Spinal N-methyl-d-aspartate receptors may contribute to the immobilizing action of isoflurane. *Anesthesia & Analgesia, 96*(1), 102-107.

Starr, A., & Farley, G. R. (1983). Middle and long latency auditory evoked potentials in cat. II. Component distributions and dependence on stimulus factors. *Hearing Research, 10*(2), 139-152.

Stoelting, R. K., & Eger, B. I. (1969). The effects of ventilation and anesthetic solubility on recovery from anesthesia: an in vivo and analog analysis before and after equilibrium. *The Journal of the American Society of Anesthesiologists*, *30*(3), 290-296.

Stolzberg, D., Wong, C., Butler, B. E., & Lomber, S. G. (2017). Catlas: an magnetic resonance imaging-based three-dimensional cortical atlas and tissue probability maps for the domestic cat (Felis catus). *Journal of Comparative Neurology*, *525*(15), 3190-3206.

Sun, Y., Lu, Y., Huang, Y., & Jiang, H. (2014). Is dexmedetomidine superior to midazolam as a premedication in children? A meta-analysis of randomized controlled trials. *Pediatric Anesthesia*, *24*(8), 863-874.

Tanaka, R., Tanaka, S., Ichino, T., Ishida, T., Fuseya, S., & Kawamata, M. (2020). Differential effects of sevoflurane and propofol on an electroretinogram and visual evoked potentials. *Journal of anesthesia*, *34*(2), 298-302.

Tenenbein, P. K., Lam, A. M., Klein, M., & Lee, L. (2006). Effects of sevoflurane and propofol on flash visual evoked potentials. *Journal of neurosurgical anesthesiology*, *18*(4), 310.

Thornton, C., Catley, D., Jordan, C., Lehane, J., Royston, D., & Jones, J. (1983). Enflurane anaesthesia causes graded changes in the brainstem and early cortical auditory evoked response in man. *British Journal of Anaesthesia*, *55*(6), 479-486.

Thornton, C., Creagh-Barry, P., Jordan, C., Luff, N., Dore, C., Henley, M., & Newton, D. (1992). SOMATOSENSORY AND AUDITORY EVOKED RESPONSES RECORDED SIMULTANEOUSLY: DIFFERENTIAL EFFECTS OF NITROUS OXIDE AND TSOFLURANE. *BJA: British Journal of Anaesthesia*, *68*(5), 508-514.

Thornton, C., Heneghan, C., James, M., & Jones, J. (1984). Effects of halothane or enflurane with controlled ventilation on auditory evoked potentials. *British Journal of Anaesthesia*, *56*(4), 315-323.

Thornton, C., Heneghan, C., Navaratnarajah, M., Bateman, P., & Jones, J. (1985). Effect of etomidate on the auditory evoked response in man. *British Journal of Anaesthesia*, *57*(6), 554-561.

Thornton, C., Heneghan, C., Navaratnarajah, M., & Jones, J. (1986). Selective effect of althesin on the auditory evoked response in man. *British Journal of Anaesthesia*, *58*(4), 422-427.

Thornton, C., Konieczko, K., Knight, A., Kaul, B., Jones, J., Dore, C., & White, D. (1989). Effect of propofol on the auditory evoked response and oesophageal contractility. *British Journal of Anaesthesia*, *63*(4), 411-417.

Vinje, M., Moe, M., Valø, E., & Berg-Johnsen, J. (2002). The effect of sevoflurane on glutamate release and uptake in rat cerebrocortical presynaptic terminals. *Acta anaesthesiologica scandinavica*, *46*(1), 103-108.

Webster, W., & Aitkin, L. (1971). Evoked potential and single unit studies of neural mechanisms underlying the effects of repetitive stimulation in the auditory pathway. *Electroencephalography and clinical neurophysiology, 31*(6), 581-592.

Wehr, M., & Zador, A. M. (2005). Synaptic mechanisms of forward suppression in rat auditory cortex. *Neuron, 47*(3), 437-445.

White, B., Abbott, L. F., & Fiser, J. (2012). Suppression of cortical neural variability is stimulus- and state-dependent. *Journal of neurophysiology, 108*(9), 2383-2392.

Wiedemayer, H., Fauser, B., Armbruster, W., Gasser, T., & Stolke, D. (2003). Visual evoked potentials for intraoperative neurophysiologic monitoring using total intravenous anesthesia. *Journal of neurosurgical anesthesiology, 15*(1), 19-24.

Woods, D. L. (1995). The component structure of the N 1 wave of the human auditory evoked potential. *Electroencephalography and Clinical Neurophysiology-Supplements Only, 44*, 102-109.

Wurtz, R. H., & Kandel, E. R. (2000). Central visual pathways. *Principles of neural science, 4*, 523-545.

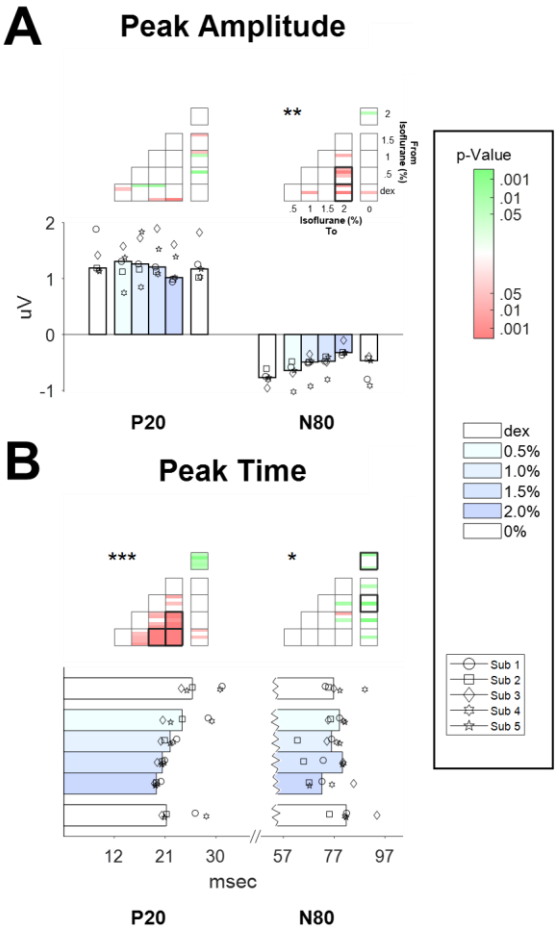
Yang, J., Isenberg, K. E., & Zorumski, C. F. (1992). Volatile anesthetics gate a chloride current in postnatal rat hippocampal neurons. *The FASEB journal, 6*(3), 914-918.

You, Y., Klistorner, A., Thie, J., & Graham, S. L. (2011). Improving reproducibility of VEP recording in rats: electrodes, stimulus source and peak analysis. *Documenta ophthalmologica, 123*(2), 109-119.

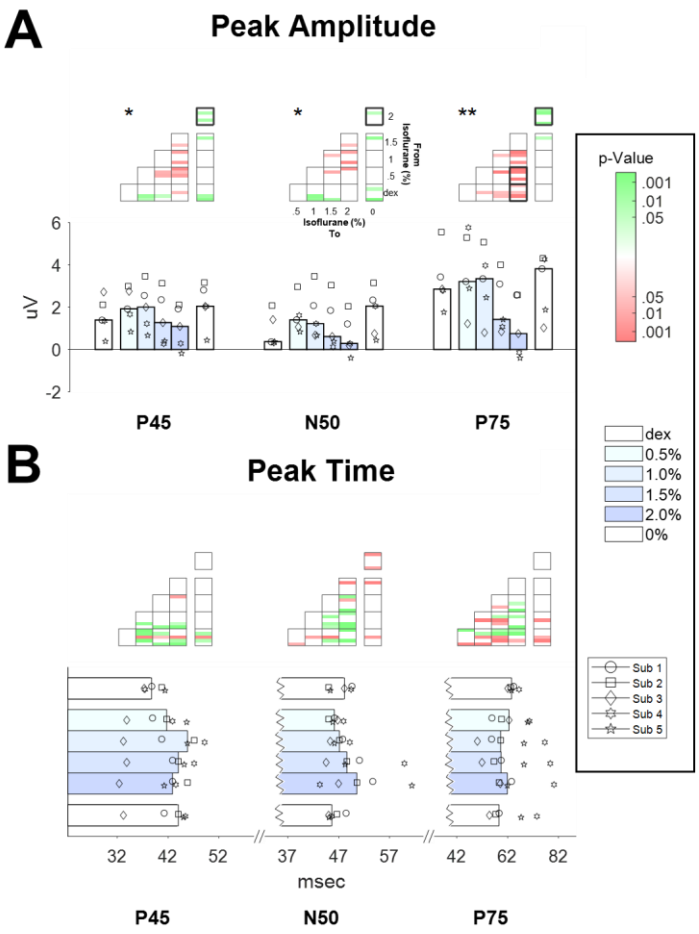
Zhang, Y., Stabernack, C., Sonner, J., Dutton, R., & Eger, E. I. (2001). Both cerebral GABAA receptors and spinal GABAA receptors modulate the capacity of isoflurane to produce immobility. *Anesthesia & Analgesia, 92*(6), 1585-1589.

Zurita, P., Villa, A., De Ribaupierre, Y., De Ribaupierre, F., & Rouiller, E. (1994). Changes of single unit activity in the cat's auditory thalamus and cortex associated to different anesthetic conditions. *Neuroscience Research, 19*(3), 303-316

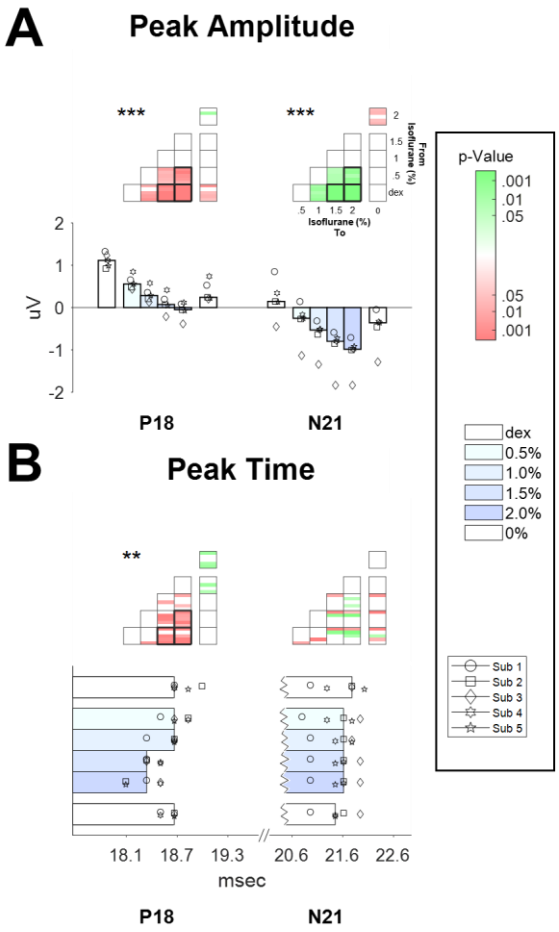
4.8. Supplementary Information



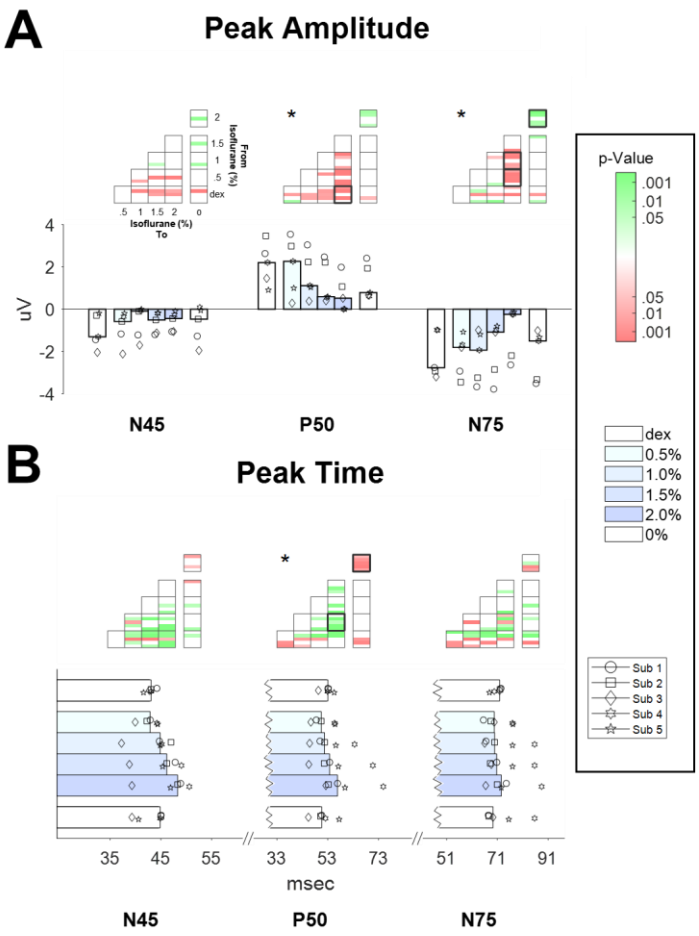
Supplementary Figure 1. Effect of isoflurane on the LL-AEP components P20 and N80. Same legends in Fig. 3 were used here.



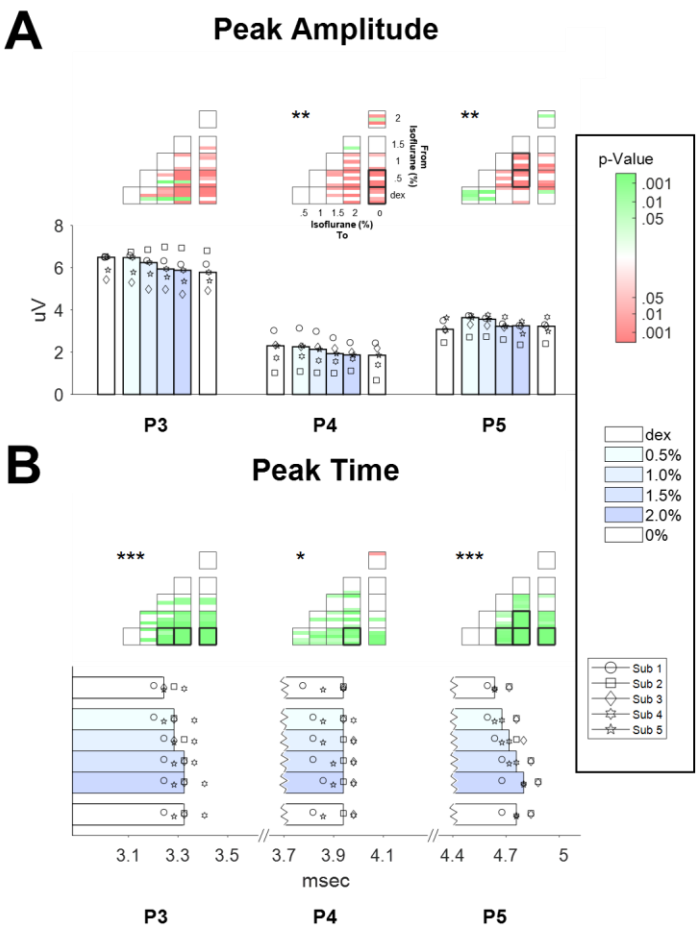
Supplementary Figure 2. Effect of isoflurane on the LL-VEP components P45, N50, and P75. Same legends in Fig. 3 were used here.



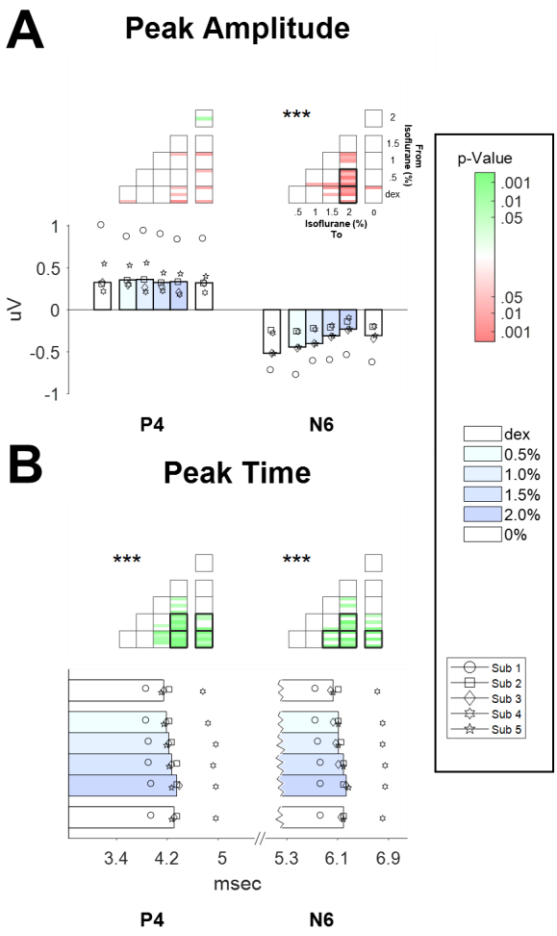
Supplementary Figure 3. Effect of isoflurane on the ML-AEP components P18 and P21. Same legends in Fig. 3 were used here.



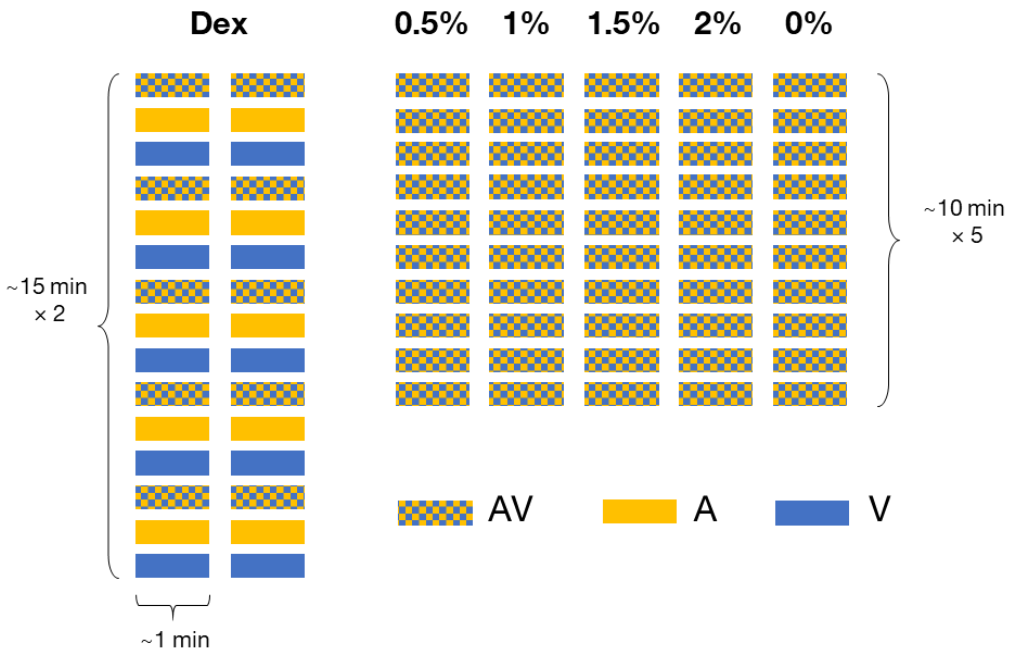
Supplementary Figure 4. Effect of isoflurane on the ML-VEP components N45, P50, and N75. Same legends in Fig. 3 were used here.



Supplementary Figure 5. Effect of isoflurane on the SL-AEP components P3, P4, and P5. Same legends in Fig. 3 were used here.



Supplementary Figure 6. Effect of isoflurane on SL-VEP components P4 and N6. Same legends in Fig. 3 were used here.



Supplementary Figure 7. Detailed breakdown recording duration.

5. Chapter 5: Temporal course of visual modulation in auditory evoked potentials in lightly-anesthetized cats

5.1. Relation to the thesis

The existing animal studies on multisensory or audiovisual interaction have been focused on one or a couple of brain areas but lacked a multi-stage, whole-brain perspective. In contrast, the existing human studies using fMRI have investigated the whole-brain involvement in audiovisual interaction and found that both cortical and subcortical brain areas are involved in audiovisual processing (see **sections 1.3.1** and **1.3.2**). There is also one human electrophysiological study (Caron-Desrochers, 2018) investigating different components of auditory evoked potentials (AEPs), including auditory brain-stem responses (ABRs), auditory middle-latency responses (MLRs), and auditory static-state evoked potentials (SSEPs), to study the effect of visual modulation along the ascending auditory pathway.

As demonstrated in Chapter 4, I have established an anesthetized/sedated cat model for auditory and visual evoked potentials. The aim of this chapter is to further expand the existing knowledge of scalp-recorded cat EEG activities in sensory processing with a focus on multisensory (audiovisual) interaction. Furthermore, I specifically chose to examine the effect of audiovisual temporal disparity (or, specifically, flash-to-click delay) on auditory evoked potentials (AEPs), to further the existing knowledge of temporal course of audiovisual interaction, the central aim of this thesis. This factor was also investigated using click-evoked local field potentials (LFPs) recorded from cat primary auditory cortex, as in Chapter 3. I hypothesized that effect of visual modulation of auditory evoked potentials (AEPs) in response to clicks was dependent on the audiovisual delay.

The data of this manuscript has been partially presented on *Association for Research in Otolaryngology (ARO) Midwinter Meeting.* The manuscript will be submitted with revision to *Scientific Report*. Modifications after the written and oral exams can be found in its published version on *Scientific Reports*.

Bao, X., Lomber, S.G. (2023). Temporal Course of Visual Modulation Revealed by Auditory Evoked Potentials in the Cat. *Association for Research in Otolaryngology (ARO) Midwinter Meeting.* (Poster SA176)

Bao, X., & Lomber, S. G. (2024). Visual modulation of auditory evoked potentials in the cat. *Scientific Reports*, 14(1), 7177.

5.2. Abstract

Visual modulation of the auditory system is not only a neural substrate for multisensory processing, but also serves as a backup inputs underlying cross-modal plasticity in deaf individuals, in which case the auditory areas that are no longer receiving acoustic inputs could be re-purposed for visual functions. Event-related potentials (ERPs) studies in human subjects have provided evidence of a multiple-stage audiovisual interactions, ranging from tens to hundreds of milliseconds after the presentation of stimuli. However, it is still unknown if the temporal course of visual modulation in the auditory ERPs can be characterized in animal models. EEG signals were recorded in a total of 14 cat subjects under dexmedetomidine sedation from subdermal needle electrodes. The auditory stimuli (clicks) and visual stimuli (LED flashes) were timed by two

independent Poison processes and were presented either simultaneously or alone. The ERPs from the visual-only condition was subtracted from that from the audiovisual condition before the comparison to the ERPs from the auditory-only condition. In the grand average across subjects, N1 amplitude showed a trend of transiting from suppression-to-facilitation with a disruption at ~100-ms flash-to-click delay. The oscillatory patterns of visual modulation in N1 amplitude also seemed to couple with several frequency bands, including 1-2 Hz, 9-10 Hz, 18-22 Hz, and 27-30 Hz. We concluded that visual modulation of auditory processing can be successfully evaluated in cats under dexmedetomidine sedation using auditory evoked potentials]. The periodic pattern can be interpreted with "phase resetting" hypothesis. Future studies can further characterize the roles played by visual modulation in the deafened and the restored auditory system.

5.3. Introduction

Brains receive sensory inputs from multiple modalities. The interaction between different sensory modalities not only affects the outcome of perception but also provide brains the opportunity of functional reorganization in the areas that loss sensory inputs under abnormal circumstances, such as blindness and deafness (Meredith, 2015). To understand the role of cross-modal interactions in brain plasticity after sensory loss and sensory restoration, it is critical to fully characterize the features of multisensory processing in adequate animal models.

It has been almost unanimously agreed that the cross-modal timing between two stimuli plays a key role in multisensory processing (Meredith, 1987, see Koelewijin for a review). An audiovisual disparity, or stimulus onset asynchrony (SOA), of ~100 milliseconds could substantially impede the perception of simultaneity (Dixon, 1980; Stevenson, 2014; Fister, 2016) and provided sufficient information for temporal order judgement (Hirsh, 1961; Vatakis, 2008).

The improvement on the performance of perception (e.g., identification reaction time or accuracy) by adding stimulus from a second modality also diminished with audiovisual SOA (Wassenhove, 2007; Harrar, 2017; Chen, 2013, 2018a, 2018b).

The range of SOAs up to 100 milliseconds, whichh cross-modal temporal processing (simultaneity and temporal order judgement) is sensitive to, has been studied in several human ERP experiments (Naue, 2011; Thorne, 2011). We referred SOAs of this range "short SOAs" and both studies have showed that short SOAs can modulate the multisensory component of ERP activities. However, longer SOAs have not been extensively studied yet.

A few literatures have shed some lights on the effect of long SOAs in multisensory processing using extracellular recording or behavior measurements. In macaque primary auditory cortex, Lakatos (2007) showed that neuronal activities evoked by click was modulated by a preceding tactile stimulus with up to 800-ms SOA. Fiebelkorn (2011) measured the fluctuated behavioral performance in detecting a near-threshold Gabor stimulus after a preceding tone beep up to a 6-second SOA. The findings in both studies have implied that the effect of long SOAs on multisensory interaction is due to the oscillations in the cortical excitability phase-locked to the preceding stimulus. Thus, we hypothesized that, in auditory ERPs, cross-modal modulation from a visual input should also occur with audiovisual temporal disparity beyond the range sensitive for multisensory temporal processing, where a periodic pattern of fluctuation may be observed.

Given that existing ERP studies on the temporal disparity of audiovisual integration provided very limited information specific to long SOAs and its spectral patterns especially in animal models (Giard & Peronnet, 1999; Molholm, 2002; Fort, 2002; Vidal, 2008; Cappe, 2010; Cappe, 2012; Mercier, 2020), the current study is aimed at investigating the association between

cat ERPs in response to auditory (click) and visual (flash) stimuli and audiovisual SOAs up to 1 second and to characterize a spectral profile of the effect of visual modulation on auditory processing. We found that the amplitude of N1 from cortical auditory evoked potentials (cAEPs) in cat under dexmedetomidine sedation was affected by audiovisual SOAs. Change in N1 amplitude as a function of SOA revealed a temporal dynamic of visual modulation in an oscillatory pattern.

5.4. Results

Cats under dexmedetomidine sedation were presented with 10 repeats of 1-min trains of clicks (auditory, A), flashes (visual, V), and unsynchronized clicks and flashes (audiovisual, AV). Then, an offline bandpass filter between 1 Hz and 10 Hz was applied for obtaining cortical auditory evoked potentials (cAEPs) and another offline bandpass filter between 10 Hz and 300 Hz was applied for obtaining middle latency responses (MLRs). Signals from V sessions were subtracted from AV sessions, generating an AV-V condition. Sweeps were extracted from AV-V and A conditions, respectively, for averaging and peak measurements (**Fig. 1A**).

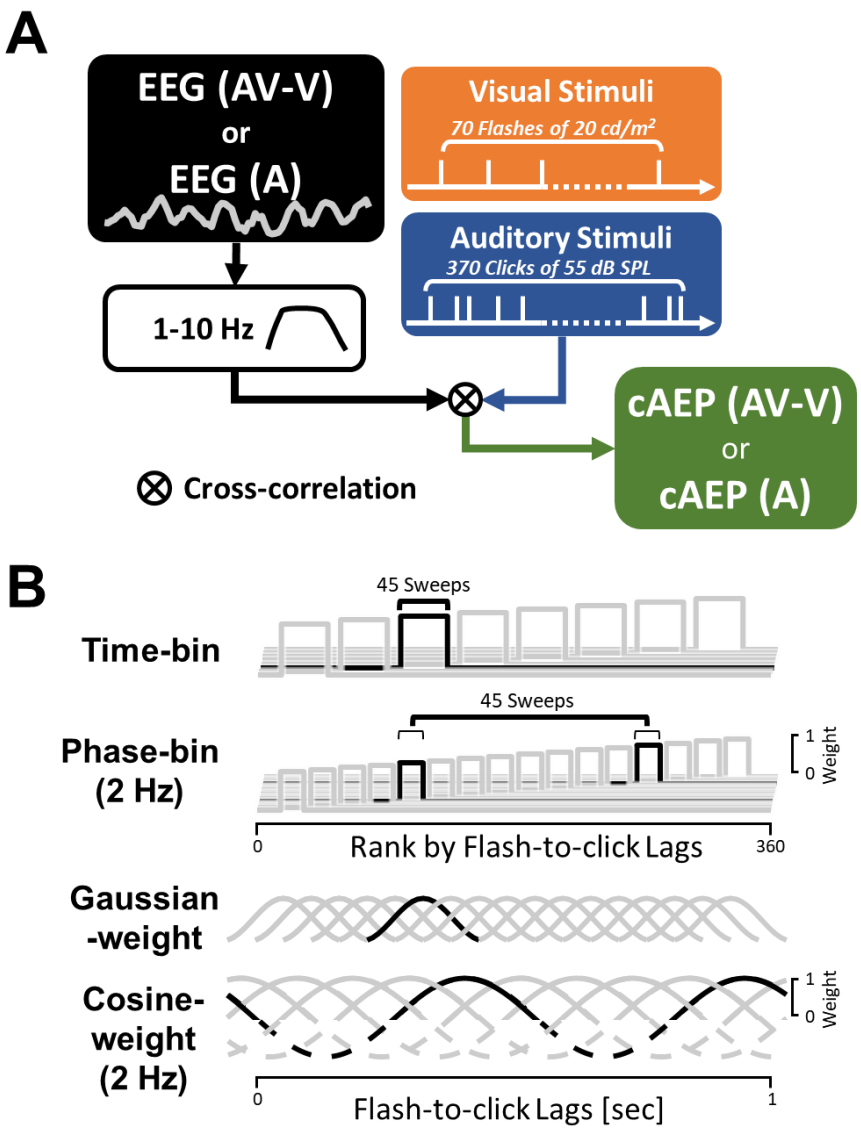


Figure 1. Recording paradigm and click-locked average of EEG signal. (A) Flow chart of EEG signal pre-processing. AV-V condition was derived from subtracting visual only (V) session from audiovisual (AV) session. EEG signals from AV-V and A conditions were filtered between 1 and 10 Hz and cross-correlated with the timestamp sequence of clicks to obtain cortical auditory evoked potentials (cAEP) regardless of audiovisual delays. **(B)** Clicks were sorted by flash-to-click lags. Only selected clicks were included into the cross-correlation process (or stimulus-locked averaging). Time-binned average, every 45 clicks (or sweeps) were included for each of 8 subgroups. Phase-binned average, same as time-binned average but with clicks sorted by phase value derived from flash-to-click lags. Gaussian-weight average, for each step of 5 millisecond flash-to-click lag, sweeps were scaled by a gain derived from a Gaussian kernel centered to the given flash-to-click lag before averaging. Three kernel widths were selected (20, 50, 100 ms). Cosine-weight average, the gain values were derived from a cosine function for a given frequency, where negative gain values were set to zero. Average with inverted (180 degrees) and orthogonal (90 and 270 degrees) phases were carried out separately. The process was repeated for two phase shifts (30 and 60 degrees) to account for the unknown optimal initial phase.

Data were collected from 14 cat subjects (mean age: 4.8 yrs, 2 males). One subject was excluded for middle latency responses (MLRs), due to a compromised recording quality at the second half of the recording period. Two additional subjects were excluded for cortical auditory evoked potentials (cAEPs). These subjects, however, were still included in the analysis of grand average waveforms.

Effect of Visual Modulation in Cat cAEPs

We first investigated the effect of visual modulation in cAEPs. From the grand-average waveforms, we observed a near-perfect overlap between the AV-V and the A conditions (**Fig. 2A**), especially for the initial 125-ms duration after click onset, suggesting a well-preserved cAEP morphology when unsynchronized visual stimuli were simultaneously present. There appeared to be an elevation of the traces starting at 150-ms after click onset in the AV-V condition. The difference between the two conditions, however, was insignificant and mostly contributed by three subjects revealed by waveforms of individual subject.

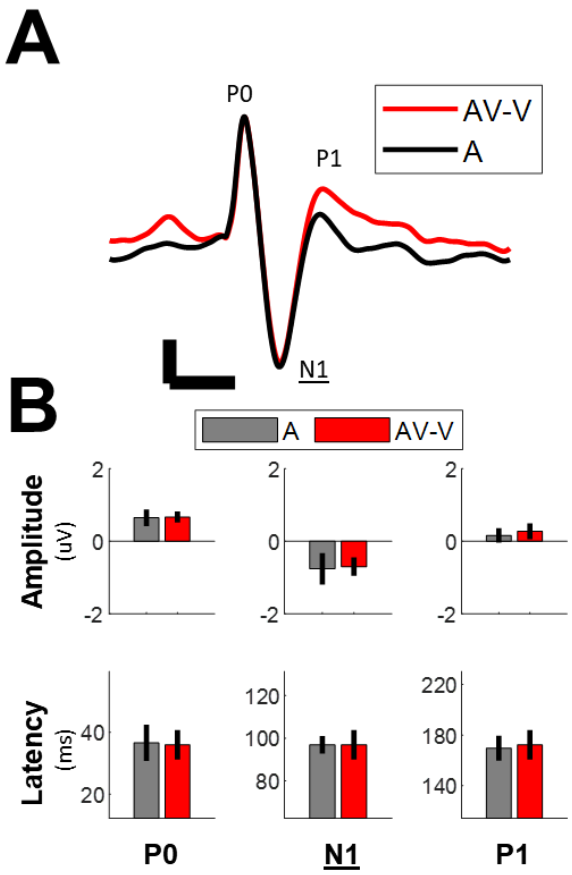


Figure 2. Cortical auditory evoked potentials (cAEPs) under unsynchronized visual modulation. **(A)** Grand average of cAEP waveform across all 14 subjects. Vertical scale, 0.25 uV. Horizontal scale, 100 ms. **(B)** Comparison of peak measurements between Auditory Only (A) and audiovisual subtracting visual (AV – V) conditions. Top, peak amplitude. Bottom, peak time. Bar, medians across subjects. Error bar, interquartile ranges.

The cAEP waveforms from the both conditions featured a prominent positive peak component about 35-ms latency, which we referred as P0, followed by a slower and wider negative peak component at about 95-ms latency post-click, which we referred as N1 (**Fig. 2A**). A second peak component, less prominent than P0, was present at about 170-ms latency, which we referred to as P1. Both peak amplitudes and latencies in these three components showed no statistically

significant difference between the two conditions across subjects (**Fig. 2B**), which is consistent with the existing knowledge that out-of-timing visual stimulus does not affect auditory processing (Caron-Desrochers, 2018).

The effect of audiovisual SOA on the effect of visual modulation In cAEPs

To examine the relationship between audiovisual temporal disparity and visual modulation of auditory processing, we performed a series of ERP averaging on subgroups of click sweeps using various strategies (**Fig. 1B**). We started by creating 8 sub-groups of 45 click sweeps ordered by the delays from their preceding flashes (i.e., flash-to-click lags) (**Fig. 3A-i**). Therefore, each of the eight sub-group cAEP waveforms represents for click responses under the visual modulation of a certain range of audiovisual temporal disparity.

We first compared the variation of peak measurements across the sub-groups. Here, we focused on N1 amplitude only. Overall, individual subjects showed a larger range of N1 amplitude across the 8 sub-groups in the AV-V condition than the A condition (**Supplementary Fig. 1B**), despite that there was a good correlation of N1 amplitude between the two conditions before the sub-grouping (**Supplementary Fig. 1A**).

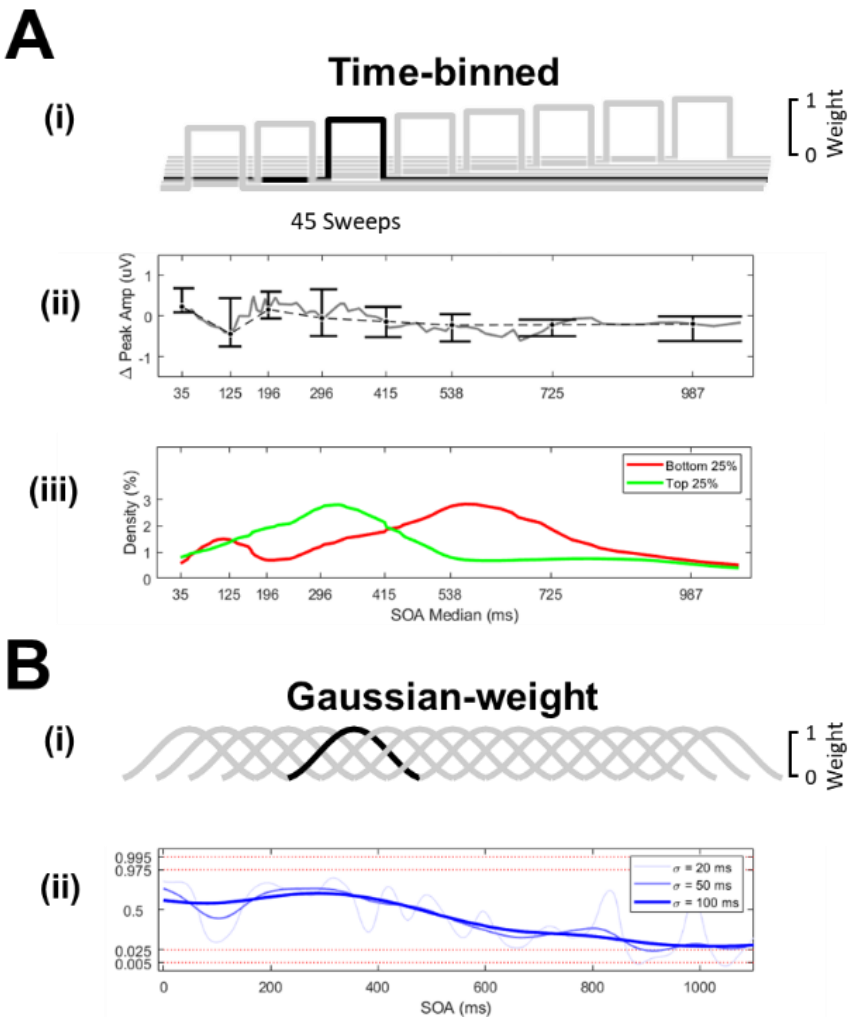


Figure 3. Effect of SOA on visual modulation in N1 amplitude. (A) The effect of visual modulation in N1 amplitude quantified by difference of N1 amplitudes between (AV – V) condition and A condition as a function of SOA time. **(B)** Kernel distribution of the top and bottom 25% of the distribution of visual modulation sampled from a finer SOA time resolution. **(C)** Temporal dynamics of visual modulation in N1 amplitude as a function of SOA generated with moving Gaussian-weight averaging process.

We performed a Friedman's test on the visual modulation of N1 amplitude and found a significant effect of SOA sub-group ($\chi^2_{Friedman}$ (df=7) = 17.697, p = 0.013 < 0.05) against the variation across subjects. Further Wilcoxon sign rank tests showed that N1 amplitude was significantly suppressed in the35-ms (p = 0.024 < 0.05) and facilitated at 987-ms group (p = 0.042 < 0.05). The rest of sub-groups failed to reveal a statistically significant visual modulation, suggesting that visual modulation in those ranges of audiovisual temporal disparity was less consistent or, more likely, the arbitrary sub-grouping was not aligned to the optimal audiovisual temporal disparity for statistical power (**Fig. 3A-ii**).

Therefore, we expanded our search for audiovisual temporal disparity of visual modulation with 66 overlapping SOA sub-groups (plotted as in gray line) (**Fig. 3A-ii**). Suppression of N1 amplitude seemed to frequently occur for SOA between 168-ms and 209-ms. The largest facilitation happened to fall between 538-ms and 725-ms SOAs. When pooled together across all subjects, the change in N1 amplitude had its top 25% values (indicating visual suppression) distributed around 320-ms SOA while its bottom 25% values (indicating visual facilitation) distributed around 560-ms SOA (**Fig. 3A-iii**).

Finally, we characterized the temporal dynamics of visual modulation in N1 amplitude using a window-weight averaging strategy (**Fig. 3B-i**). Briefly, for any given SOA, sweeps were averaged with weight values derived from a Gaussian kernel centered at this SOA. The width of the Gaussian kernels was controlled by the parameter σ, which was selected to be 20-, 50-, or 100-ms concerning the trade-off between SOA resolution and the signal-to-noise ratio of averaged waveforms. We observed an overall transition from visual suppression to facilitation of N1 amplitude at ~500-ms SOA, as well as an early and transient facilitation at ~100-ms SOA (**Fig. 3B-ii**). Such temporal dynamic was partially captured by the sub-grouping strategy and tested

across individual subjects. In addition, the use of smaller σ values revealed a substantial fluctuation that appears to conform to some periodic patterns. This observation was consistent with previous findings in audio-tactile interaction of macaque primary auditory cortex (Lakatos, 2007) and audio-visual interaction of human visual cortex (Romei, 2012), which motivated us to continue characterizing the spectral profile of the visual modulation in N1 amplitude next.

Patterns of periodicity in visual modulation locked to flash onset

One underlying mechanism of cross-modal modulation involves oscillation of cortical activities. The ongoing cortical oscillation in one sensory area can be modulated by activity from another brain area (Keitel, 2017) or phase-reset by an external event (VanRullen & Macdonald, 2012). It was reported that stimulus from a non-auditory modality could modulate click-evoked responses in the primary auditory cortex through resetting the phase of cortical oscillation (Lakatos, 2007; Kayser, 2008).

To reveal such periodic patterns time locked to the flash onset in visual modulation of N1 amplitude, we selected 6 frequency values, ranging logarithmically from 2 Hz to 64 Hz. For each frequency, SOA phase values were derived from SOA time values before being designated to one of the eight 45-degree-wide sub-groups (**Fig. 4A-i**). Again, we first compared the range of N1 amplitudes across the phase subgroups between the AV-V and the A conditions. We found a more apparent visual modulation in terms of the range for phase-grouping than time-grouping (**Supplementary Fig. 1C**). Then, we contrasted the two conditions and compared the range of difference in N1 amplitude across the phase sub-groups among the six frequencies, approximating the range of difference as an estimation of the magnitude of the periodic variation (or equivalently

the consistency of phase-locking) of visual modulation in each individual subject. The smallest and the largest variation averaged across subjects were found at 4 Hz and 16 Hz, respectively (**Fig. 4A-ii**).

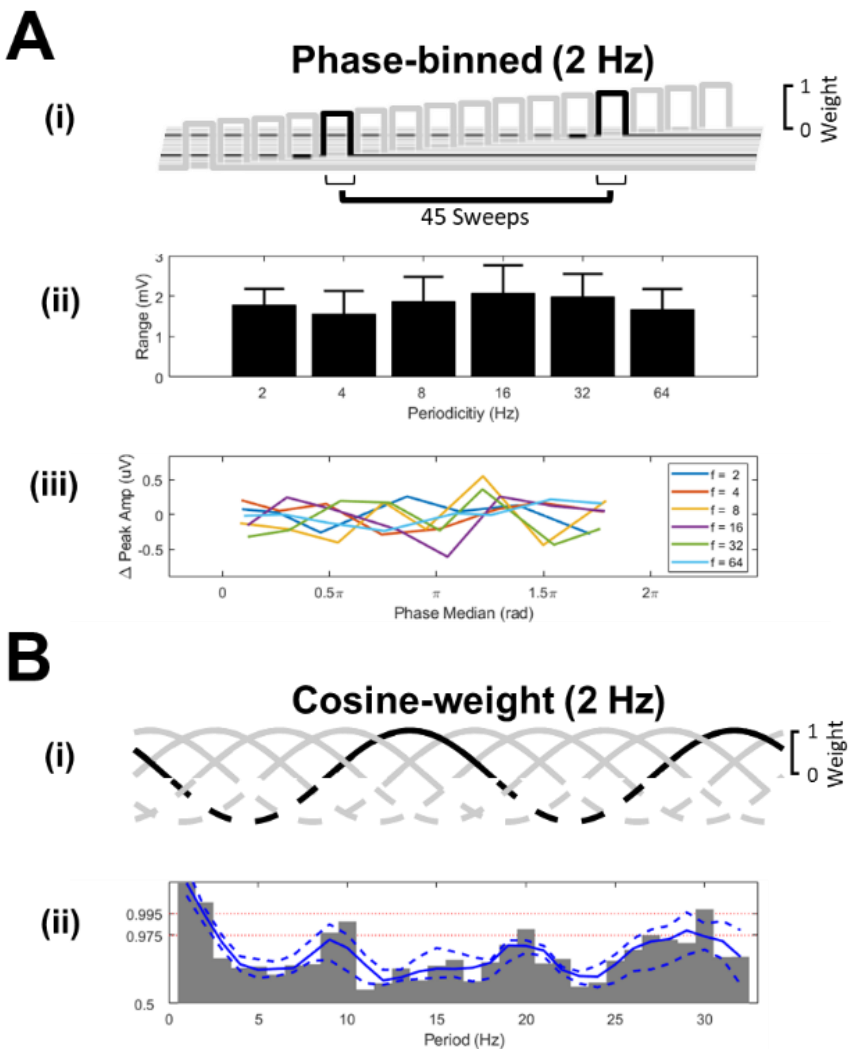


Figure 4. SOA periodicity of visual modulation in N1 amplitude. (A) Visual modulation in N1 amplitude as a function of SOA phase for different frequencies. **(B)** Range of visual modulation across phase bins in each frequency. Error bar, interquartile range across subjects. **(C)** Spectrum of the visual modulation dynamics generated with shifting Cosine-weight averaging process.

To examine the effect of frequency on the consistency of phase locking across subjects, we performed a two-way ANOVA (6 frequencies × 8 phase sub-groups) on the N1 amplitude differences. However, there was no significant interaction between the two factors (**Fig. 4A-iii**). We then performed six one-way ANOVA tests for the effect of phase sub-groups by each frequency. We found a significant main effect of phase sub-group ($F_{7,70} = 3.09$, $p = 0.007 < 0.01$) only at 8 Hz, which suggested a periodic interval of ~125-ms SOA between two adjacent cycles of the visual modulation.

Ultimately, we characterized a spectral profile of visual modulation in N1 amplitude from the grand average across subjects (**Fig. 4B-ii**) using a cosine-weight averaging approach. In the range from 1 Hz to 32 Hz, the spectral magnitude peaked beyond our significance criteria at ~1.5 Hz (delta-to-theta band), ~10 Hz (alpha band), ~20 Hz (beta band), and ~30 Hz (beta band). The observation of an unevenly distributed spectral pattern in the temporal dynamics of visual modulation provided further support to the phase-reset hypothesis proposed for visual modulation in the primary auditory cortex (Lakatos, Gross, & Thut, 2019). In addition, our data suggested that the influence of 10-Hz oscillation in EEG initiated by visual stimuli (VanRullen & Macdonald, 2012) could be beyond visual cortex and have a modulatory effect on a non-visual function, following 10-Hz periodicity pattern or its harmonics.

Effect of Visual Modulation in Cat MLRs

For comparison, we also measured P20 amplitude from middle latency responses (MLRs) and performed the same analysis. We found that, in the temporal domain, the effect of SOA time-bin was not significant in P20 amplitude ($\chi^2_{Friedman}$ (df=7) = 13.49, $p = 0.061$), however, Wilcoxon sign rank tests revealed significant facilitation at 124-ms SOA ($p = 0.014 < 0.05$) and 194-ms SOA

($p = 0.032 < 0.05$) (**Fig. 5A-i**). The temporal dynamics of P20 amplitude demonstrated similar fluctuations as in the visual modulation of N1 amplitude from cAEP but not the overall transition from visual suppression to facilitation (**Fig. 5A-ii**). We also observed less periodicity in the visual modulation of P20 amplitude, from both phase sub-group averaging and cosine-weight averaging. On the spectral profile, high magnitudes were noticed around 26 Hz, corresponding to beta-band EEG (**Fig. 5B**). There was not a high magnitude specific to delta-to-theta frequencies as was observed in the visual modulation of N1 amplitude and the other cAEP components (**Supplementary Fig. 2**).

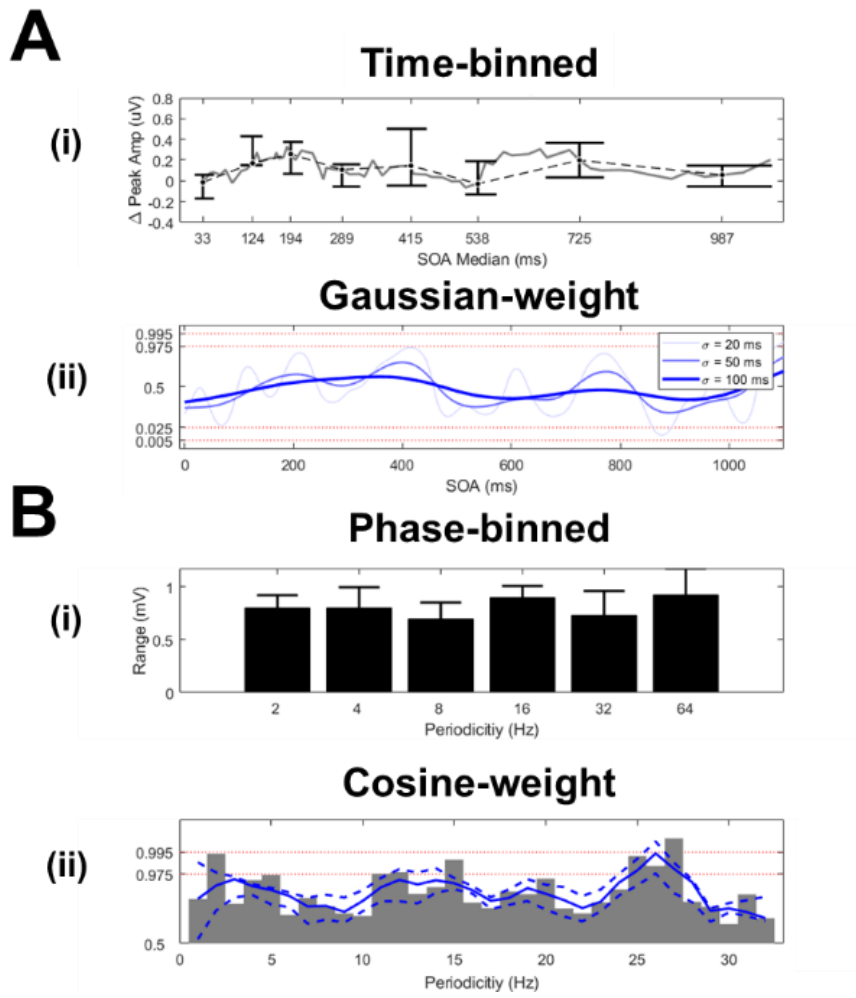


Figure 5. Effect of SOA on visual modulation in MLR P20 amplitude. (A) Visual modulation in N1 amplitude as a function of SOA phase for different frequencies. **(B)** Range of visual modulation across phase bins in each frequency. Error bar, interquartile range across subjects. **(C)** Spectrum of the visual modulation dynamics generated with shifting Cosine-weight averaging process.

5.5. Discussions

Cross-modal modulation of auditory processing beyond temporal binding window

In this study, we examined and demonstrated the effect of audiovisual temporal disparity or stimulus onset asynchrony (SOA) on visual modulation of cortical auditory evoked potentials (cAEPs). The audiovisual interaction was investigated using similar approaches in two previous

human ERP studies, with SOAs below 100 milliseconds (Thorne, 2011) and 70 milliseconds (Naue, 2011), respectively. A few extracellular recordings studies examined SOAs up to 500 milliseconds in the superior colliculus (Meredith, 1987) and 320 milliseconds in auditory cortices (Kayser, 2008). These studies have made the discoveries of the neural correlates to the "temporal window of integration" that were measured behaviorally, demonstrating strong evidence for a "coincidence detector" as a neurophysiological mechanism (Leone, 2013; Rowland, 2007; Truszkowski, 2017).

Long SOAs, despite of not likely being involved with the temporal integration or temporal processing (perception of multisensory simultaneity and temporal order), are still possible for effective cross-modal modulation of sensory processing. This idea has been supported by both behavior data (Fiebelkorn, 2011; Diederich, 2012; Diederich, 2014) and some neurophysiological evidence (Lakatos, 2007; Romei, 2012). Lakatos (2007) pointed out that the optimal SOAs for tactile modulation of sound-evoked neuronal activities in their data were corresponded to the periodic intervals of several EEG oscillations. According to the "phase reset" hypothesis they proposed, a preceding tactile stimulus resets the phase of ongoing neural oscillations in the primary auditory cortex, which in turn determines the state of fluctuating auditory excitability. When the SOA between the preceding tactile stimulus and the following auditory stimulus is aligned to the high-excitability, up-phase of neural oscillation, the auditory stimulus evokes a larger response than when tactile-auditory SOA is aligned to the low-excitability, low-phase of neural oscillation. The observation of excitability fluctuation has been further evidenced with various behavioral and electrophysiological measurements, including extracellular recording (Kayser, 2008), human ERP (Naue, 2011), phosphine induced by transcranial magnetic stimulation (Romei, 2012; Romei, 2013), and reaction time (Diederich, 2012, 2014; Fiebelkorn, 2011).

While the "phase-reset" hypothesis has served as a fundamental mechanism in many extended scenarios of sensory and perceptual processing, explaining the coupling/entraining between continuous stimulus signals (e.g., pulse train, narrowband noise, and speech) (Luo, 2007; Atilgan, 2018; Lakatos, 2008) and neural activities or between neural activities from different regions (Keitel, 2017), its neural mechanism is still not fully understood. One of many questions that are still unclear is whether the spectral distribution carried by stimulus has a direct association with the spectral distribution of the elevated EEG oscillation. In macaque primary auditory cortex (Lakatos, 2008), 1.5-Hz pulse trains entrained neural activities of delta band (1-4 Hz) in the cortical regions of task-relevant modality, in which case the spectral distribution in stimuli matched that of entrained neural signals. By contrast, in awake ferret auditory cortex (Atilgan, 2018), narrowband noise with coherent audiovisual envelope at about 7 Hz showed enhanced local field potential between 10.5 and 12.5Hz, which belongs to alpha band (8-12 Hz), in which case the spectral distributions were unmatched.

Here we propose that a spectrally white stimulus sequence, such like the Poisson train of flashes, can be a useful tool to investigate the oscillation of neural modulation driven but not "dominated" by a cross-modal stimulus input. In our results, it appeared that, although weakly, the fluctuation of excitability participated in the cross-modal visual modulation of auditory activities, even when neither auditory nor visual stimuli featured an entraining frequency. Future studies can incorporate both white and non-white stimulus signals to interrogate the origin of the neural oscillations involved in the various functions of perceptions and cognitions.

Critical role of EEG/ERP studies in multisensory research

One of the many missions of the future multisensory research is to converge the knowledge established from extracellular recordings in animal models and from whole-brain imaging in human. While data of intracranial recordings in human are still rare and challenging to obtain, scalp-EEG recordings from large animal models are quickly developing as a uniquely useful neurophysiological approach (marmoset: Itoh, 1015, 2021; cat, Presacco 2018a, 2018b; Richardson, 2022; Heidari, 2019; Mitzelfelt, 2023).

Electrical and magnetic mappings of whole-brain activities during audiovisual perception have provided valuable insights on its neural mechanism involving intra-cortical functional connectivity (Keitel, 2017) and topographic re-distribution (Cappe, 2010). Human auditory evoked potentials have been well-characterized for a variety of components as neural correlates to sound processing at different stages of ascending auditory pathway (Scherg, 1989; review: Eggermont, 2002). The current study is by far the first scalp-recorded EEG multisensory study in animal models, and is, infrequently in literature, focused on auditory evoked potentials under visual modulation. We compared ERPs from the auditory-only condition with a derived condition by subtracting signal of the visual-only condition from the audiovisual condition, rather than compare the difference between audiovisual condition with a derived condition by "sum of the auditory and the visual conditions". This allowed us to select peak components time-locked to auditory stimuli, which are supposed to have better interpretability for auditory processing.

In summary, we reported N1 amplitude in scalp-recorded auditory evoked potentials (AEPs) from cats under dexmedetomidine sedation as a sensitive measurement for cross-modal visual modulation of auditory processing. The delay function, sampled with both sparse grouping approach and fine-resolution weight-average approach, revealed a short-SOA effect peaking at ~100 ms as well as a long-SOA effect characterizing the time course of visual modulation over

~1-second period. Informed by the spectrum estimated with a SOA-derived phase-locking averaging approach, we found evidence in support of periodic patterns (~1.5 Hz, ~10 Hz, ~20 Hz, and ~30 Hz) presenting in the temporal course of visual modulation driven by spectrally white stimulus input signals. With the advantages of our animal models and experiment paradigms, future studies are expected to characterize the spectrotemporal features in normal and sensory-deprived subjects and to identify the neural mechanism underlying cross-modal interactions.

5.6. Methods

All procedures were conducted in compliance with the National Research Council's Guide for the Care and Use of Laboratory Animals (8th edition; 2011) and the Canadian Council on Animal Care's Guide to the Care and Use of Experimental Animals (1993. Furthermore, the following procedures were also approved by Animal Care Committee (DOWB) for the Faculty of Medicine and Health Sciences at McGill University.

Animal preparation and anesthesia protocol

After subjects were sedated using dexmedetomidine (0.04 mg/kg, Dexdomitor, Zoetis) injected intramuscularly, the left eye was occluded using a black contact lens so that visual stimuli were presented unilaterally. Phenylephrine (Mydfrin, Alcon) was applied to the right eye to dilate the pupil, and saline drops were used as lubrication. Subjects were placed on a water-circulated heating pad (TP-400, Gaymar). Once vital signs (heart rate and SpO2) were stable, two 15-minute recording sessions were carried out while the subject was breathing 80% oxygen (Dispomed). At the end of the two recording sessions, data collection terminated in nine subjects and continued in the other five under isoflurane anesthesia for a separate study. Subject’s vital signs and electrode impedance were checked between the two sessions. At the end of data collection, electrodes and

contact lens were removed before atipamezole (Antisedan, Zoetis) was administrated intramuscularly to facilitate recovery from the dexmedetomidine sedation.

Visual and auditory stimuli

The visual stimuli consisted of flashes were presented to subjects from a 5-mm-diameter light-emitting diode (~11 degrees of visual field, LED, DigiKey). The intensity of flash stimuli was calibrated to 10 lux by adjusting the voltage magnitude of a 300-us-long squared pulse as the input signal to the LED. The auditory stimuli were 300-us-long clicks emitted by an 8-cm-diamter loudspeaker (Fostex). The sound level of the click stimuli was calibrated to 55 dB SPL using a sound meter (Model 2250, B&K). Both auditory and visual stimulation signals were generated by the same digital-to-analogue processor (RZ2, TDT). The LED was attached to the top of the loudspeaker and placed 8-cm away from the subject at the direction of 45 degrees right to the midline.

To manipulate the timing of auditory and visual stimulus, two independent, 57-second-long pulse trains for triggering clicks and flashes, respectively, were pre-made in Matlab using a Poisson random process and loaded into the stimulus/recording software (Synapse, TDT). The auditory stimulus train contained 370 click pulses and the visual stimulus train contained 70 flash pulses. The inter-pulse dead time was set to 20 milliseconds and 500 milliseconds for the auditory and visual stimuli, respectively. The auditory and the visual stimulus trains, lasting ~1 minutes, always started and stopped simultaneously in each session. Auditory only (A), visual only (V), and audiovisual (AV) stimulus trains were played alternatively in order and repeated for 10 times.

EEG recording and signal processing

Three 27G stainless steel needles were placed subdermal as recording electrodes. The active electrode was placed near the midpoint of subject's interaural line, while the reference electrode was placed beneath the right ear (ipsilateral to the side of visual stimulation). The ground electrode was placed on the subject's dorsum (~10 cm behind should blade near the midline). The impedance of both active and reference electrodes was maintained below 3 kOhm during recording. The signal was amplified and digitized with a pre-amplifier (RA4LI/RA4PA, TDT), streamed onto a digital signal processor (RZ2, TDT), and stored on a computer hard drive.

All data analysis was performed offline. Signal was digitally notched at 30 Hz, 60 Hz, 120 Hz, 180 Hz, and 240 Hz before passing through band-pass filters (1-10 Hz for cortical auditory evoked potentials and 10-300 Hz for middle latency responses). Then, signals from same stimulus conditions were averaged. For AV-V condition, AEPs were derived from subtracting visual only (V) session average from audiovisual (AV) session average. For A condition, AEPs was derived from auditory only (A) session average.

Flash-to-click lags were calculated for each individual click as the delay of its onset to the onset of its preceding flash for audiovisual stimulus. Epochs were extracted between 50-ms pre-click and 250-ms post-click for cortical auditory evoked potentials and between 20-ms pre-click and 100-ms post-click for middle latency responses.

In time-binned sub-group averaging, sweeps were ordered ascendingly by flash-to-click lags. Bins were created for every 45 sweeps and labeled as stimulus onset asynchrony (SOA) of the median flash-to-click lags. The first 360 sweeps were included, with the remaining 10 sweeps discarded. In phase-binned sub-group averaging, 6 frequencies (2, 4, 8, 16, 32, 64 Hz) were selected. For different frequencies, phase values of flash-to-click lags were calculated as the

remainder of dividing each lag by the invert of the frequency. Sweeps were then ordered by the phase values and binned-averaged similarly as in time-bin averaging.

In Gaussian-weight averaging, SOAs were selected from 0 to 1000-ms with a 5-ms step. For each SOA, a Gaussian kernel function with one of the three widths (σ = 20, 50, 100 msec) was centered at the SOA. Clicks within $\pm 3\ \sigma$ range were included into the average with weight values given by the Gaussian kernel functions. Click sweeps with flash-to-click lags more deviating away from the SOA or the peak of Gaussian kernel therefore contributed less to the averaged waveform. In cosine-weight average, the Gaussian kernel functions were replaced by cosine functions with its negative values set to zero. Similar to Gaussian-weight average, for each frequency from 1 to 32 Hz at 1-Hz step, sweeps were averaged with weights given by the cosine function output at their click onset. Additional averaged waveforms were also obtained for reversed initial phase (180 degree), as well as their orthogonal rotations (90 and 270 degrees). The difference value of the peak measurements between the two phase-reserved averaged waveforms was combined into its orthogonal version using quadratic mean, which gives an estimate of the magnitude of oscillation at the given frequency. Additional estimates were obtained with two shifts of initial phase (30 and 60 degrees).

Extraction of N1 amplitude

From the averaged waveform, peaks were detected by local maxima algorithm and compete for the best fit to a template. The peak(s) with closest peak time to the template was considered first. In the case of the same peak-time distance from the two sides, the waveform was interpolated to a doubled time resolution iteratively until only one final peak time was found.

Statistics

For time-binned and phase-binned sub-group averaging, N1 amplitudes were derived from each individual subject. Friedman test, Wilcoxon test, and two-way ANOVA test were performed on Matlab using Statistics and Machine Learning Toolbox™. P-values smaller than 0.05 were reported as statistically significant.

For Gaussian-weight and cosine-weight averaging, N1 amplitudes were derived using the grand average across all subjects. To test for statistical significance, 1000 permutations were performed by randomizing the mapping between the sweeps and their flash-to-click lags, which allows the measures of visual modulation derived from the original data normalized to a value between 0 and 1.

5.7. References

Atilgan, H., Town, S. M., Wood, K. C., Jones, G. P., Maddox, R. K., Lee, A. K., & Bizley, J. K. (2018). Integration of visual information in auditory cortex promotes auditory scene analysis through multisensory binding. Neuron, 97(3), 640-655.

Cappe, C., Thut, G., Romei, V., & Murray, M. M. (2010). Auditory–visual multisensory interactions in humans: timing, topography, directionality, and sources. Journal of Neuroscience, 30(38), 12572-12580.

Chen, Y. C., & Spence, C. (2013). The time-course of the cross-modal semantic modulation of visual picture processing by naturalistic sounds and spoken words. Multisensory Research, 26(4), 371-386.

Chen, Y. C., & Spence, C. (2018). Audiovisual semantic interactions between linguistic and nonlinguistic stimuli: The time-courses and categorical specificity. Journal of Experimental Psychology: Human Perception and Performance, 44(10), 1488.

Chen, Y. C., & Spence, C. (2018). Dissociating the time courses of the cross-modal semantic priming effects elicited by naturalistic sounds and spoken words. Psychonomic Bulletin & Review, 25, 1138-1146.

Diederich, A. et al. (2012) Saccadic reaction times to audiovisual stimuli show effects of oscillatory phase reset. PLoS ONE 7, e44910

Diederich, A., & Colonius, H. (2015). The time window of multisensory integration: relating reaction times and judgments of temporal order. Psychological review, 122(2), 232.

Diederich, A., Schomburg, A., & Van Vugt, M. (2014). Fronto-central theta oscillations are related to oscillations in saccadic response times (SRT): an EEG and behavioral data analysis. PLoS One, 9(11), e112974.

Dixon, N. F., & Spitz, L. (1980). The detection of auditory visual desynchrony. Perception, 9(6), 719-721.

Eggermont, J. J., & Ponton, C. W. (2002). The neurophysiology of auditory perception: from single units to evoked potentials. Audiology and Neurotology, 7(2), 71-99.

Fiebelkorn, I. C., Foxe, J. J., Butler, J. S., Mercier, M. R., Snyder, A. C., & Molholm, S. (2011). Ready, set, reset: stimulus-locked periodicity in behavioral performance demonstrates the consequences of cross-sensory phase reset. Journal of Neuroscience, 31(27), 9971-9981.

Harrar, V., Harris, L. R., & Spence, C. (2017). Multisensory integration is independent of perceived simultaneity. Experimental Brain Research, 235, 763-775.

Hirsh, I. J., & Sherrick Jr, C. E. (1961). Perceived order in different sense modalities. Journal of experimental psychology, 62(5), 423.

Kayser, C., Petkov, C. I., & Logothetis, N. K. (2008). Visual modulation of neurons in auditory cortex. Cerebral Cortex, 18(7), 1560-1574.

Keitel, A., Ince, R. A., Gross, J., & Kayser, C. (2017). Auditory cortical delta-entrainment interacts with oscillatory power in multiple fronto-parietal networks. Neuroimage, 147, 32-42.

Keitel, A., Ince, R. A., Gross, J., & Kayser, C. (2017). Auditory cortical delta-entrainment interacts with oscillatory power in multiple fronto-parietal networks. Neuroimage, 147, 32-42.

Lakatos, P., Chen, C. M., O'Connell, M. N., Mills, A., & Schroeder, C. E. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. Neuron, 53(2), 279-292.

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. science, 320(5872), 110-113.

Lakatos, P., Gross, J., & Thut, G. (2019). A new unifying account of the roles of neuronal entrainment. *Current biology, 29*(18), R890-R905.

Leone, L. M., & McCourt, M. E. (2013). The roles of physical and physiological simultaneity in audiovisual multisensory facilitation. i-Perception, 4(4), 213-228.

Luo, H., & Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. Neuron, 54(6), 1001-1010.

Meredith, M. A., & Allman, B. L. (2015). Single-unit analysis of somatosensory processing in the core auditory cortex of hearing ferrets. European Journal of Neuroscience, 41(5), 686-698.

Meredith, M. A., Nemitz, J. W., & Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. Journal of Neuroscience, 7(10), 3215-3229.

Naue, N., Rach, S., Strüber, D., Huster, R. J., Zaehle, T., Körner, U., & Herrmann, C. S. (2011). Auditory event-related response in visual cortex modulates subsequent visual responses in humans. Journal of Neuroscience, 31(21), 7729-7736.

Naue, N., Rach, S., Strüber, D., Huster, R. J., Zaehle, T., Körner, U., & Herrmann, C. S. (2011). Auditory event-related response in visual cortex modulates subsequent visual responses in humans. Journal of Neuroscience, 31(21), 7729-7736.

Romei, V., Gross, J., & Thut, G. (2012). Sounds reset rhythms of visual cortex and corresponding human visual perception. Current biology, 22(9), 807-813.

Romei, V., Murray, M. M., Cappe, C., & Thut, G. (2013). The contributions of sensory dominance and attentional bias to cross-modal enhancement of visual cortex excitability. Journal of Cognitive Neuroscience, 25(7), 1122-1135.

Rowland, B. A., Stanford, T. R., & Stein, B. E. (2007). A model of the neural mechanisms underlying multisensory integration in the superior colliculus. Perception, 36(10), 1431-1443.

Scherg, M., Vajsar, J., & Picton, T. W. (1989). A source analysis of the late human auditory evoked potentials. Journal of cognitive neuroscience, 1(4), 336-355.

Thorne, J. D., De Vos, M., Viola, F. C., & Debener, S. (2011). Cross-modal phase reset predicts auditory task performance in humans. Journal of Neuroscience, 31(10), 3853-3861.

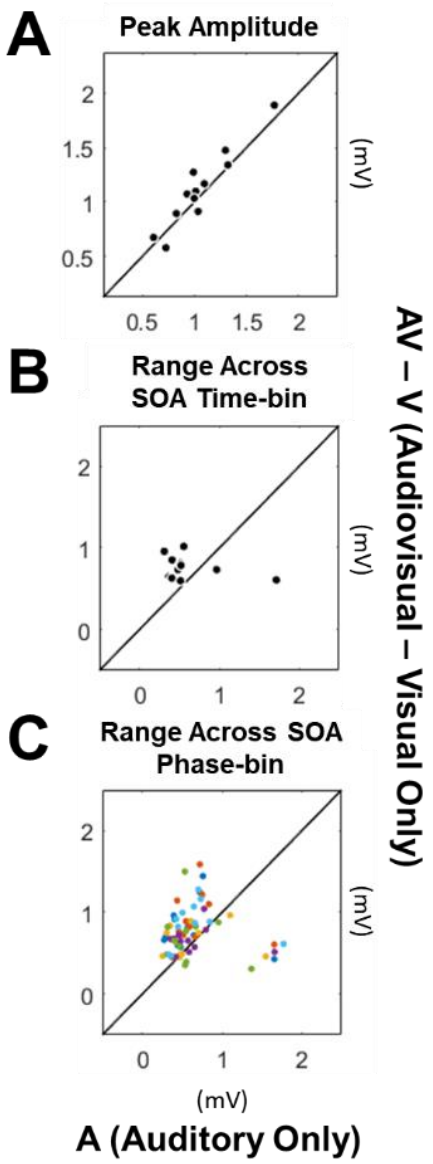
Thorne, J. D., De Vos, M., Viola, F. C., & Debener, S. (2011). Cross-modal phase reset predicts auditory task performance in humans. Journal of Neuroscience, 31(10), 3853-3861.

Truszkowski, T. L., Carrillo, O. A., Bleier, J., Ramirez-Vizcarrondo, C. M., Felch, D. L., McQuillan, M., ... & Aizenman, C. D. (2017). A cellular mechanism for inverse effectiveness in multisensory integration. Elife, 6, e25392.

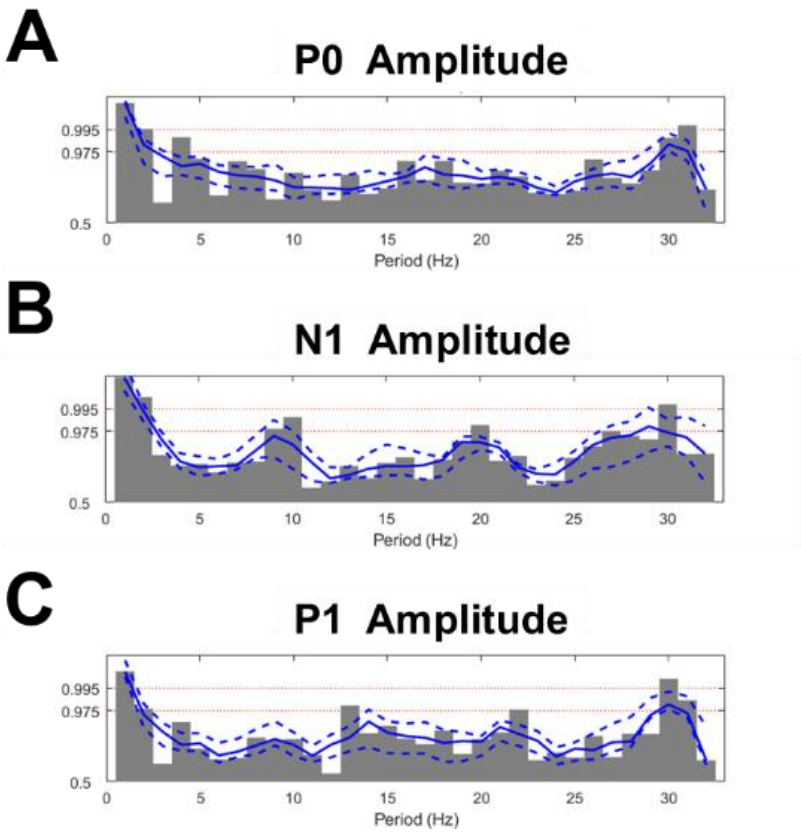
Van Wassenhove, V., Grant, K. W., & Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. Neuropsychologia, 45(3), 598-607.

VanRullen, R., & Macdonald, J. S. (2012). Perceptual echoes at 10 Hz in the human brain. *Current biology, 22*(11), 995-999.

5.8. Supplementary Information



Supplementary Figure 1. N1 amplitude and its range under visual modulation. Scatter plots for N1 amplitude (AV – V) condition versus A condition with dots indicating individual subjects. (A) N1 amplitude from the averaged waveforms across SOA. (B) Range of N1 amplitudes across 8 SOA time-bins in each subject. (C) Range of N1 amplitudes across 8 SOA phase-bins for each of 6 frequencies (color-coded) in each subject.



Supplementary Figure 2. SOA periodicity in visual modulation in three cAEP peak components. Spectrum of the visual modulation dynamics generated with shifting Cosine-weight averaging process for P0 amplitude (A), N1 amplitude (B), and P1 amplitude (C).

6. Chapter 6: General Discussion

6.1. Overview

In this thesis, I revisited the visual modulation of auditory responses using cat as an animal model and investigated the effect of audiovisual temporal disparity invovled, using awake extracellular recording and lightly anesthetized EEG recording as experiment approaches. Most previous studies reported single temporal integration window, that is, audiovisual interaction was optimized for an audiovisual delay of less than 100-msec audiovisual delay (Bremen, Massoudi, Van Wanrooij, & Van Opstal, 2017; Kayser, Petkov, & Logothetis, 2008; Schormans et al., 2017). In my thesis, however, the results derived from both experimental approaches demonstrated multiple and complex temporal integration windows, i.e., audiovisual delay from 100 to more than 1000 milliseconds.

In this chapter, I will first review the implications of this finding regarding the existing hypotheses/explanations proposed by previous multisensory research. Next, I will relate the current findings and the neurophysiology of visual projections to the auditory cortex, in the context of cross-modal interaction and cross-modal plasticity. Thirdly, I will go over some issues regarding the selection of auditory stimulus (which was clicks) and recording techniques (which were both field potentials). Finally, I will briefly discuss some potential directions for future studies, with a focus on cat model and its advantages.

6.2. Implications on the temporal perspective of audiovisual processing

The effect of audiovisual temporal disparity on visual modulation of auditory responses that was characterized in both dexmedetomidine sedated cats and awake cats indicated that there

exist multiple windows or intervals of optimal delays for visual facilitation or suppression. The observation of multiple temporal integration windows in the primary auditory cortex was also reported in one electrophysiological study in macaque A1 (Lakatos, Chen, O'Connell, Mills, & Schroeder, 2007) using audio-tactile stimuli, and the authors proposed that the long-delay audio-tactile interactions were due to the phase reset of the ongoing EEG oscillations and the cortical excitability by the preceding tactile stimulus.

6.2.1. Optimal delay for audiovisual interaction: multiple or single integration window?

A couple of previous studies have characterized the neurophysiological activation of audiovisual stimuli as a function of audiovisual temporal disparity (Kayser et al., 2008; Schormans et al., 2017). In these studies, the ranges of audiovisual delays were limited to a couple hundreds of milliseconds and only a single temporal integration window was found at population level (Kayser et al., 2008; Schormans et al., 2017). In macaque primary auditory cortex, the study showed an optimal flash-to-noise delay of 20-, 40-, and 80-msec for visual suppression of auditory responses. The other study, in rat lateral extrastriate visual cortex, demonstrated that a comparable flash-to-noise delay (20-, 40- or 60-msec) was optimal for multisensory enhancement. The range of delays in both studies were in accordance with the critical delays for temporal order and simultaneity judgement observed behaviorally (Harrar, Harris, & Spence, 2017; Schormans et al., 2017; Van der Burg & Goodbourn, 2015). In comparison, we did find the largest attenuation of click LFPs by visual stimulus at flip-to-click delays less than 100-msec from 15% ~ 20% of the recording electrodes in both subjects (Chapter 3), which is consistent with macaque data (Kayser et al., 2008). However, we also demonstrated visual enhancement at flip-to-click delays much greater than 100-msec for a large number of electrodes. These findings were not implausible, since

it has been already previously reported that audiovisual Ventriloquist effect still occurred for temporal delay of 200 milliseconds (McGovern, Roudaia, Newell, & Roach, 2016), suggesting the presence of audiovisual interaction with large temporal disparity.

It has also been noted in the literature that extended temporal integration window could be observed in patients with dyslexia (Hairston, Burdette, Flowers, Wood, & Wallace, 2005) and autism spectrum disorders (ASDs) (Foss-Feig et al., 2010). In addition, compared to adult participants, teenagers not only had larger temporal binding window (Hillock‐Dunn & Wallace, 2012) but also demonstrated a larger shift in point of subjective simultaneity (PSS) after being exposed to asynchronous stimuli (Noel, De Niear, Van der Burg, & Wallace, 2016). These findings appear to suggest that a broad temporal integration window at young age or in patients is supporsed to be narrowed during the process of normal brain maturation. The pruning or erasing process of existing functionality is likely associated with the development of GABAergic inhibitory system (Berardi, Pizzorusso, & Maffei, 2000; Dorrn, Yuan, Barker, Schreiner, & Froemke, 2010; Rubenstein & Merzenich, 2003; Zheng & Knudsen, 1999). Therefore, further investigation on the contribution of cortical excitation and inhibition may help the understanding of how temporal integration windows are constructed and modified through development.

6.2.2. Effect of temporal disparity: temporal and non-temporal tasks

Aside from directly asking participants to make explicit judgement on simultaneity or temporal order of audiovisuals stimuli, the effect of temporal disparity was also revealed in the performance of various perception tasks, including a detection task (J. Miller, 1986), a discrimination task (Van der Burg, Olivers, Bronkhorst, & Theeuwes, 2008), and a recognition

task (Y. Yuan, Wayland, & Oh, 2020). It was shown that the effect of both audiovisual improvement and disruption depended on audiovisual delay (Harrar et al., 2017; Kawase et al., 2016; Mégevand, Molholm, Nayak, & Foxe, 2013). For example, McGurk effect, where subjects misheard /ba/ as /da/ when viewing a speaker articulating /ga/, was much weaker when audiovisual delay exceeded 100 msec (Kawase et al., 2016).

The relationship between the temporal (simultaneity/temporal order judgement) and non-temporal processing (e.g., McGurk effect) was discussed by a few studies where the both were tested in the same group of participants (Harrar et al., 2017; Mégevand et al., 2013). Both studies demonstrated smaller inter-subject variation in optimal audiovisual delay for the multisensory benefit in a perception task (i.e., a decrease in reaction time) than inter-subject variation of temporal binding window derived from temporal processing tasks. Another study showed that the width of temporal binding window is negatively correlated with the occurrence of McGurk effect (Stevenson, Zemtsov, & Wallace, 2012), suggesting a close dependency between audiovisual perception and audiovisual temporal processing. In support of this argument, it was found that selective TMS stimulation at the superior temporal sulcus (STS), a critical brain region associated to the perception of temporal asynchrony, also disrupted McGurk effect (Beauchamp, Nath, & Pasalar, 2010).

Cognitive and physiological mechanisms underlying the visual modulation of auditory responses and hearing behaviors have been proposed (Stekelenburg & Vroomen, 2007; Van der Burg et al., 2008), in the names of warning or preparatory effect (Bertelson, 1967; Los & Van Den Heuvel, 2001), expectancy effect (Niemi & Näätänen, 1981), alertness effect (Posner & Boies, 1971), cross-modal cueing effect (McDonald, Teder-SaÈlejaÈrvi, & Hillyard, 2000), sensory gating effect (Lebib, Papo, de Bode, & Baudonnière, 2003; Oray, Lu, & Dawson, 2002;

Stekelenburg & Vroomen, 2007), and binding effect (Van der Burg et al., 2008). In these studies, a delay of no more than 1 seconds between visual and auditory stimuli was often adopted. However, these mechanisms are often difficult to be experimentally differentiated between each other at behavioral levels (Van der Burg et al., 2008).

The temporal disparity in audiovisual stimuli can be extended into temporal coherence under certain circumstances. Recent studies using speech sentences as stimuli have shown that artificial visual stimulus that tracking the envelop of speech sentences increased the percentage of word recognition (Yi Yuan, Lleo, Daniel, White, & Oh, 2021; Y. Yuan et al., 2020) while there was no visual benefit when the temporal dynamics of the visual stimuli were uncorrelated to the speech envelop. Similarly, in an auditory localization task, the visual benefit was dependent on the correlation between the auditory and visual pulse trains (Parise, Harrar, Ernst, & Spence, 2013; Parise, Spence, & Ernst, 2012). EEG signal was also demonstrated to be phase-locked better to prolonged audiovisual stimuli when there was a temporal coherence in auditory and visual modalities (Crosse, Di Liberto, & Lalor, 2016; Nozaradan, Peretz, & Mouraux, 2012). What's important to this thesis is that the performance of spatial localization tolerates a 66-msec systematic audiovisual delay (Parise et al., 2012). Several studies have shown that the delay in temporally coherent audiovisual stimuli was autonomously incorporated by the sensory system (Navarra, Fernández-Prieto, & Garcia-Morera, 2013; Navarra et al., 2005; Vatakis, Navarra, Soto-Faraco, & Spence, 2007), which will be discussed in the next section.

6.2.3. Adaptation to asynchronized audiovisual stimuli

It is also possible that the interaction with long audiovisual disparity that we observed is not recruited for multisensory integration but for multisensory adaption. Studies have shown that the optimal delay for the perception of audiovisual simultaneity can be shifted by the exposure to asynchronous audiovisual speech (Navarra et al., 2005; Vatakis et al., 2007) and non-speech stimuli (Fujisaki, Shimojo, Kashino, & Nishida, 2004; Navarra et al., 2013; Noel et al., 2016). When subjects were exposed to auditory-leading stimulus, the perceived subjective simultaneity (PSS) of visual-to-auditory delay shifts negatively. Oppositely, when subjects were exposed to visual-leading stimulus, the PPS of visual-to-auditory delay shifts positively (Navarra et al., 2005; Vatakis et al., 2007). However, one study found that exposure to asynchronous audiovisual stimuli with a delay greater than 700-msec induced a shift of PPS only in visual-leading condition but not auditory-leading condition (Navarra et al., 2013).

Related to the exposure effect reported in those literature, the late temporal integration windows in A1 neurons we observed may be involved in detecting the cross-correlation of asynchronous audiovisual stimuli and trigger the short-term plasticity that underlies the behavioral shifting of PPS. Furthermore, presuming the role of the auditory cortex in multisensory adaption, it is also possible that the neural circuits in the auditory cortex is distinct from the visual cortex, as suggested by the asymmetric finding of visual- but not auditory-leading adaption (Navarra et al., 2013). In order to fully understand the interplay of temporal disparity and temporal coherence in audiovisual processing, the stimuli used in the current thesis may be further adapted to incorporate zero and non-zero optimal lags for varying cross-correlations in future studies.

6.2.4. Cross-modal phase-reset hypothesis

When multiple temporal integration windows were first discovered in macaque auditory cortex, it was explained with a "phase-reset" hypothesis by the authors (Lakatos et al., 2007). It was suggested that a preceding tactile stimulus reset the phase of ongoing EEG oscillation at different frequency bands (Lakatos et al., 2007). Accordingly, when one EEG oscillation cycles back from out-phase to in-phase, the excitability of neurons of the auditory cortex is also switched from down-state to up-state. The intervals of one cycle for different EEG bands are also different (i.e., the invert of frequency), corresponding to different temporal integration windows.

In support of the "phase-reset" hypothesis, the authors later found that attention allocated to one sensory modality could promote stimuli in the attended modality in phase-resetting the ongoing local field potentials in primary cortical areas for the ignored modality (Lakatos, Karmos, Mehta, Ulbert, & Schroeder, 2008). In macaque auditory cortex, the visually evoked LFP was enhanced when subjects attending visual but not auditory stimuli. Vice versa, auditory evoked LFPs in the visual cortex was enhanced when subjects attending auditory but not visual stimuli.

To summarize this section, the effect of temporal disparity on audiovisual interaction can be interpreted in a variety of circumstances, including audiovisual temporal processing (6.2.1), audiovisual facilitation/suppression (6.2.2), adaption to audiovisual asynchrony (6.2.3), and cross-modal phase-reset (6.2.4). The capability of neurons in the auditory cortex integrating a preceding non-auditory stimulus that is temporally separated for hundreds of milliseconds from an auditory stimulus can be closely associated with the behavioral outcomes in those circumstances.

6.3. The anatomic foundations of audiovisual processing

6.3.1. Source of visual afferents in cat auditory cortex

This thesis mainly studied the functional perspectives of audiovisual processing, with no intention on making any extrapolation on the anatomic foundations, which is a complex question by itself and should be addressed with retrograde tracer and histological staining. Retrograde tracer deposited in A1 identified more than 10% of labeled neurons projecting from the anterolateral and the posterolateral lateral suprasylvian sulcus (ALLS and PLLS, respectively), both of which were considered visual areas (Chabot, Butler, & Lomber, 2015). Neurons in both areas are responsive to stimulus located in a large azimuth range of visual fields but only within 20-40 degree away from the horizontal meridian plane (Palmer, Rosenquist, & Tusa, 1978). It was shown that about 20% of neurons in PLLS were selective to one of the five modes of motion tested, and a majority of them (70%) were selective to a radiation motion either expanding or contracting (Li, Li, Chen, Wang, & Diao, 2000). It is unknown how these neurons would respond to checkerboard stimulus.

In addition, the same retrograde tracer study identified more than 30% of labeled neurons projecting from a variety of other cortical auditory areas, especially from the dorsal zone of auditory cortex (DZ), the anterior auditory field (AAF), and the secondary auditory cortex (A2) (Chabot et al., 2015). By analyzing the laminar distribution of cell bodys, it was inferred that projections from AAF and DZ were mainly feedforward, while projections from A2 were for feedback modulation (Hackett, 2011). All these three areas were found to receive projections from visual areas. A recently published studied reported ~75% of visually responsive neurons in cat DZ (Merrikhi, Kok, Carrasco, Meredith, & Lomber, 2022). In fact, there were almost same proportion of neurons labeled by retrograde tracer deposited in DZ were from visual cortical areas as from auditory cortical areas, and those from ALLS and PLLS altogether account for more than 30% of

the total labeled neurons (Kok, Chabot, & Lomber, 2014). On the other hand, 47% of neurons in AAF were converted to be visually responsive following deafness (Meredith & Lomber, 2011), which was accompanied by a significant increase of tracer-labeled projections from visual areas especially from ALLS and anterior ectosylvian visual area (AEV). Those findings suggested the existence of subthreshold visual inputs into hearing AAF. Therefore, both AAF and DZ could be a good candidate as a relaying station underlying the visual modulation from visual areas (i.e., ALLS, PLLS, AEV) in the primary auditory cortex. In comparison, A2 is a less likely source for such second-order visual modulation, as there were less than 10% of tracer-labeled neurons projecting from visual areas, most of which were from AEV (Butler, de la Rua, Ward-Able, & Lomber, 2018).

In short, visual afferents from visual areas (ALLS, PLLS, and AEV) to A1, both directly and indirectly (through AAF and DZ), are good candidates for further investigations to identify the functional sources of visual modulation in A1 neurons.

6.3.2. Synaptogenesis following deafness

Recent studies have found that synapses in the supragranular layers, where most visual projections to the auditory cortex arrive, were characterized by increased spine density and spine diameter following deafness (Clemo, Lomber, & Meredith, 2016, 2017). These findings have led to a new hypothesis, that is, instead of relying on new afferents from a visual area growing inward to the auditory cortex, the functional plasticity following deafness is achieved by pre-existing visual afferents generating new synapses on to auditory neurons. This could explain why the observed difference in visual afferents using retrograde tracer between hearing and deaf cats (Kok

et al., 2014) could not fully account for the functional plasticity observed behaviorally (Lomber, Meredith, & Kral, 2010). In another word, there may exist equivalent number of visual afferents in hearing as in deaf auditory cortex, despite of less or weaker synaptic connections.

Since such visual afferents that exist in hearing auditory cortex do not fully express their potentials in making synaptic connections yet, it can be speculated that these visual afferents may not directly drive the firing of auditory neurons but rather modulate their auditory responses (Ghazanfar, Maier, Hoffman, & Logothetis, 2005; Kayser et al., 2008; M. A. Meredith & B. L. Allman, 2015). Therefore, to understand the visual processing in the auditory cortex, interrogating the visual modulation of auditory responses and measuring local field potentials should be more sensitive than examining spikes elicited by visual stimulus alone. In the next section, I will discuss the rationale of selecting click as acoustic stimuli and the limitation of field potentials as a measure of neural activation.

6.4. Auditory evoked potentials and local field potentials elicited by clicks

6.4.1. Click stimulus as opposed to its several alternatives (noise, tone, vocalization, and speech)

A variety of types of acoustic stimuli have been used in characterizing the effect of audiovisual temporal disparity. These stimuli include click paired with flash (Parise et al., 2013), noise paired with flash (Meredith, Nemitz, & Stein, 1987), audiovisual clips of tool-using (Stevenson & Wallace, 2013), vocalization paired with mouth movement (Ghazanfar et al., 2005; Kayser et al., 2008), or speech paired with shape with its size modulated by speech envelope (Yi Yuan et al., 2021).

In macaque primary auditory cortex, two previous studies both used vocalization stimuli (Ghazanfar et al., 2005; Kayser et al., 2008). Given the special role of vocalization in social communication, it is possible that the processing of these stimuli involves feedback modulation from higher order cortical neurons, such as those reported in anterior supratemporal plane (Perrodin, Kayser, Logothetis, & Petkov, 2015) and anterior medial and anterior fundus face patches (Khandhadia, Murphy, Romanski, Bizley, & Leopold, 2021; Perrodin et al., 2015). Only one of the two macaque studies systematically investigated the role of temporal disparity (Kayser et al., 2008), using a pair of noise and flash stimuli with varying delays. This stimulus paradigm was also tested in anesthetized cats (Meredith et al., 1987) and ferrets (Bizley, Nodal, Bajo, Nelken, & King, 2007), but all these studies reported a pattern of single temporal integration window.

In contrast, multiple integration windows were revealed in one previous study, which characterized the modulation of click evoked responses by a preceding tactile stimuli in macaque primary auditory cortex (Lakatos et al., 2007), not with noise stimulus. Therefore, it would be interesting for future studies to re-investigate the multiple temporal integration windows with expanded repertoire of auditory and visual stimuli.

Stimulus-specificity in audiovisual processing has been reported in previous research. Temporal processing in human was found to depend on the types of stimulus (Stevenson & Wallace, 2013). Temporal binding window measured by a simultaneity judgement task was broader for speech stimuli than non-speech stimuli (e.g., noise/flash). Similarly, in macaque auditory cortex, an artificial shape tracking the movement of mouth as visual stimuli was less effective in modulating auditory responses to vocalization (Ghazanfar et al., 2005). However, human speech and animal vocalization are more complex acoustic stimuli than tone/noise burst

and click, not only in the perspective of their ecological relevance but also their temporal structures. Such direct comparison between different stimulus types in audiovisual processing is scarce. Using narrowband noise, one recent study found an effect of temporal congruency on visual modulation of auditory responses in anesthetized ferrets (Atilgan et al., 2018). A narrowband noise is enriched with temporal structure similar to those observed in speech despite of the lack of meaning. Therefore, this study suggested that the processing of temporal congruency of audiovisual stimuli may arise from the primary auditory cortex before higher order cortical areas are recruited.

Most existing studies investigating how the primary auditory cortex is involved in visual modulation used broadband (white) noise stimulus (Bizley et al., 2007; Kayser et al., 2008; M Alex Meredith & Brian L Allman, 2015; Meredith & Lomber, 2011). Auditory neurons in non-primary auditory cortex demonstrated complex frequency selectivity (Phillips & Orman, 1984; M. L. Sutter & Schreiner, 1991), which makes noise stimulus a more effective option in terms of activating these neurons. In the primary auditory cortex, however, click responses had the shortest response latency in compared to tone and noise stimulus (Camalier, D'Angelo, Sterbing-D'Angelo, de la Mothe, & Hackett, 2012). It was also found in kitten that with increasing age A1 neurons became more responsive to click trains but not to amplitude-modulated noises (Eggermont, 1993). Therefore, the functional networks activated by click versus noise stimuli may be intrinsically different. This may seem to be trivia but another reason of choosing click stimuli as acoustic stimulus for eliciting local field potentials and auditory evoked potentials in the current thesis is based off its signal property. A click stimulus has the shortest possible duration and a relatively broad spectral distribution of energy.

Although click stimulus has the best temporal precision, for neurons residing in the primary auditory cortex, a pure tone stimulus may activate a more restricted portion of thalamocortical projections, due to the point-to-point frequency tonotopic organization (L. M. Miller, Escabı, Read, & Schreiner, 2001; Winer, Miller, Lee, & Schreiner, 2005) along the auditory ascending pathways. It has been shown that 17 out of 61 neurons in A1 responsive to pure tones of varying frequency and intensity did not respond to broad-band noise (Phillips, Orman, Musicant, & Wilson, 1985), suggesting that pure tone stimulus, with properly selected frequency and intensity, could be more effective than broadband noise in activating A1 neurons.

Using a two-tone stimulus paradigm, it has been revealed that a preceding pure tone suppressed the responses to a second pure tone, regardless of whether the two tones had the same or different frequencies, which suggested a synergistic contribution from co-existing excitatory and inhibitory inputs in neuronal firing pattern (M. Sutter, Schreiner, McLean, O'connor, & Loftus, 1999; Valentine & Eggermont, 2004). Consistent with two-tone suppression, it has also been shown that increasing stimulus density decreased response gain of dynamic random chord (DRC) stimuli, which is a stochastic train of tone pips with varying frequency (Rabinowitz, Willmore, Schnupp, & King, 2011; Valentine & Eggermont, 2004). Although it is unknown whether neuronal responses to stochastic clicks is also modulated by stimulus density (i.e. rate of click) in a similar way, it was shown by periodic click trains with varying click rate that smaller intervals between clicks would more likely induce response suppressions.

The incorporation of continuous click pulses with stochastic inter-click intervals as acoustic stimulus in the current thesis is one of its major distinction from previous research of similar topics. Unlike a single pair of auditory and non-auditory stimulus pulses, the novel stimulus paradigm is aimed at probing how cat A1 neurons respond to sound contrasting its acoustic context

(Barbour & Wang, 2003). Using the same type of stimuli, it has been shown that auditory neurons rapidly adapted to acoustic background and became less responsive to clicks (Gourévitch & Eggermont, 2008). The neural circuits underlying such context-based adaption of neuronal responses may also contribute to the visual modulation of auditory activities that we observed.

In short, there is a limited extent to which the findings of this thesis can be generalized to non-click stimuli, without further testing of pure tone, noise and speech stimuli of various types. This may be further complicated by the variant use of anesthetics reported in the literature, as there was discussion on the hidden features of the auditory cortex under anesthesia (Wang, 2007).

6.4.2. Auditory evoked potentials (AEPs) and local field potentials (LFPs) elicited by clicks

In human, with the introduction of event-related potential (ERPs) to multisensory research, there is an ample number of studies reporting ERP components specific to multisensory processing (Caron-Desrochers, Schönwiesner, Focke, & Lehmann, 2018; Fort, Delpuech, Pernier, & Giard, 2002; Giard & Peronnet, 1999; Musacchia, Sams, Nicol, & Kraus, 2006; Nozaradan et al., 2012; Stekelenburg & Vroomen, 2007). In one of the pioneering studies, participants were asked to discriminate between one object (540 Hz beep with horizontal ellipse) and the other (560 Hz beep with vertical ellipse) (Giard & Peronnet, 1999). Participant's reaction time was faster for audiovisual objects than auditory-only or visual-only objects. In addition, electrode sites near auditory areas also showed an amplified activity 90~110 ms after stimulus onset, a time window corresponding to auditory N1 component. In a later study, however, a visual suppression was observed in N1 elicited by speech audiovisual stimuli (Stekelenburg & Vroomen, 2007). The

authors argued that the effect was unique to speech processing and therefore was not observed using simple stimuli.

It is hard to compare the results between the two studies, as the spatial resolution provided by scalp-recorded EEG signal does not a allow for differentiation between human primary auditory cortex (i.e., Heschl's gyrus) and higher order auditory areas specific for speech processing (e.g., superior temporal gyrus). It is still challenging to accurately estimate the source of neural generator of scalp-recorded EEG (Olejniczak, 2006), even with the use of high density electrode configuration. For the same reason, the possibility cannot be ruled out that a non-auditory brain area (e.g., the visual cortex) may contribute to the audiovisual interaction in auditory N1 component reported in Chapter 5.

Local field potentials (LFPs) have much improved spatial specificity than scalp EEGs. However, LFPs recorded from electrodes located inside cat A1 may still pick up signal from neighbouring areas, which is termed as volume conduction (Bremen et al., 2017; Fernández-Ruiz et al., 2013; Herreras, 2016). It was estimated that LFPs could spread to recording sites more than one centimeter away in the direction perpendicular to the cortical surface (Kajikawa & Schroeder, 2011). Across cortical laminae, the component contributed by distant sites in LFP signals might be reduced by calculating the second-order spatial derivatives of LFPs, which is also known as current-source density (Kajikawa & Schroeder, 2011; Lakatos et al., 2007; Schormans et al., 2017).

Although there is no doubt that ECoG and sEEG recordings have absolute advantage over scalp EEG in determining the source of neural generator, both of them can only be performed in patients based on their clinical needs (e.g., treatment for epilepsy) due to the invasiveness of the techniques. Therefore, they are not expected to be used expensively in hearing loss population,

which is larger and more heterogenous than neurosurgical patients. Therefore, existing non-invasive electrophysiological techniques (e.g., scalp recorded EEG) still remain to be the primary approach among electrophysiological techniques of various types in order to ultimately understand cross-modal interaction in hearing and deaf human and to serve hearing restoration. Functional MRI, a non-electrophysiological technique, is known for the finest spatial resolution among all non-invasive neurophysiological measurements, which is still being improved for increasing intensity of magnetic field and other aspects (Uğurbil, 2021). For example, it can be used in participants with normal hearing or deaf participants before cochlear implant but , unfortunately, not those already wearing cochlear implant, due to the risk cast by the magnetic field on the metal part of implant (Bawazeer, Vuong, Riehm, Veillon, & Charpiot, 2019).

Alternative neural imaging techniques, such as functional near-infrared spectroscopy (fNIR), does not involve a magnetic field and therefore can be used in studying cross-modal interaction of current cochlear implant (CI) users (Anderson, Lazard, & Hartley, 2017). As was shown by a fNIR study, the increase of visual activation in the auditory cortex from pre- to post-implant measurements was positively correlated to the measurements of speech processing (Anderson, Wiggins, Kitterick, & Hartley, 2017). It was also shown that the difference between auditory activation in the visual cortex and visual activation in the auditory cortex in CI users was positively correlated with performance of speech recognition (Chen, Sandmann, Thorne, Bleichner, & Debener, 2016).

In summary, for the interpretation of neural generator to the signal of both AEPs and LFPs reported in Chapter 3 and 5, the effect of volume conduction should be 227onsidered. Localizing the source of change in the neural activities in cochlear implant recipients is critical for differentiating cross-modal and intra-modal plasticity and remains a challenge in the longitude

assessment of sensory restoration in cochlear implant recipients. More research is needed, maybe in combined with other neural imaging techniques, such as fMRI and fNIR, to identify the neural generators of AEPs and LFPs involved in visual modulation.

6.5. Awake and anesthetized cat as an animal model for studying cross-modal interaction

6.5.1. Cat as a unique animal model for audiovisual interaction

The advantages of using cat as an animal model for studying cross-modal interaction lies on several aspects. First of all, previous cat studies have charted the anatomy and function in a variety of cortical and subcortical brain areas for unisensory and multisensory processing by electrophysiological studies (Meredith et al., 2011; Merrikhi et al., 2022; Merrikhi, Kok, Lomber, & Meredith, 2023) and retrograde tracer studies (Butler, Chabot, & Lomber, 2016; Butler, de la Rua, et al., 2018; Butler, Sunstrum, & Lomber, 2018; Chabot et al., 2015; Kok et al., 2014; Wong, Chabot, Kok, & Lomber, 2015). In particular, hierarchical processing of auditory inputs to cat auditory cortex has been compared with that in non-human primates (Lee & Winer, 2008; Schreiner & Winer, 2007). These literatures have laid the foundations for the incorporation of new knowledges such as the findings of this thesis.

Secondly, the size of cat brains is particularly suitable for a reversible deactivation technique with cryoloop (Lomber, Payne, & Horel, 1999). This technique has been used in cats and monkeys (Takei, Lomber, Cook, & Scott, 2021; Wong, Pearson, & Lomber, 2018) to allow one or multiple areas of cortex to be deactivated while subject's behavior is being assessed. It has several advantages compared to reversible chemical deactivation, including better control of the duration of inactivation and faster recovery process (Lomber, 1999). However, the nature of its

working mechanism poses an inevitable limit on the minimal effective area (10–75 mm^3), which compromises the specificity of deactivation especially when a smaller animal model (i.e., rodent) is used.

Finally, the duration of pre-adult development in cats is much longer than rodents. It takes more than one year of time for a newborn kitten to fully develop into an adult cat (as in 16-year-old human) (HOYUMPA et al., 2010), which provides sufficient time for age-specific experiment procedures or treatments. In comparison, the equivalent period of development in rodents has a duration of only 7 weeks (Barrow & Schmitt, 2017). It will become more challenging in rodents when treatment in an earlier development stage (e.g., infancy) is aimed at. On the other end, cats also have a longer life expectancy (13-14 years) than rodents (3 years), which may be even further increased with veterinary science that is being continuously advanced in clinical setting (Cozzi, Ballarin, Mantovani, & Rota, 2017). At the same time, cats give multiple birth like other smaller mammals, which led to a relative lower cost for breeding as a research animal than non-human primates (Nardone et al., 2017; Tajiri et al., 2013). In short, the life cycle and reproduction mode of cats make them a very useful animal model for studying plasticity following sensory impairment or deprivation.

Finally, the application of functional MRI in cats made it a better animal model for studying sensory neuroscience (T. A. Brown et al., 2014; T. A. Brown et al., 2013; Butler, Hall, & Lomber, 2015; Hall et al., 2014; Hall, Butler, & Lomber, 2016; Hall & Lomber, 2015; Stolzberg, Butler, & Lomber, 2018). Several techniques related to cat fMRI have been developed, including anesthesia protocol (Levine, Li, Barnes, Lomber, & Butler, 2020), three-dimensional cortical atlas and tissue probability (Stolzberg, Wong, Butler, & Lomber, 2017), as well as versatile sampling strategies (Hall et al., 2014).

6.5.2. Effect of anesthetics

In the chapter 4 of this thesis, the effect of isoflurane on auditory/visually evoked potentials was examined. Although surveys showed that sevoflurane was becoming more commonly preferred by anesthetists in recent years, isoflurane had been widely used in clinical practice for >30 years since its approvalby the Food and Drug Administration (FDA) in 1979. For more than 30 years, a vast amount of research on isoflurane has been carried out in human, in vivo animal models, and in vitro cultured cells. The mechanism of isoflurane as an anesthetic, and as a confounding factor in neurophysiological experiments using anesthetized animals as well, is still being investigated today.

Type-A γ-aminobutyric acid (GABA) receptors are deemed as the primary target of isoflurane, like many other intravenous and volatile anesthetics. At low concentrations, isoflurane is known to potentiate GABA-induced chloride current (Harrison, Kugler, Jones, Greenblatt, & Pritchett, 1993; Krasowski et al., 1998; Nakahiro, Yeh, Brunner, & Narahashi, 1989; Zhang, Stabernack, Sonner, Dutton, & Eger, 2001). It was shown that isoflurane, under specific conditions, could also directly induce hyperpolarizing chloride current through GABA receptors (Jia et al., 2008; Yang, Isenberg, & Zorumski, 1992). Either direct or indirect facilitation of GABAergic inhibition caused by isoflurane may serve as the major underlying mechanism of its effect on AEPs and VEPs.

The effect of GABAergic inhibition on auditory responses has been shown along rat ascending auditory pathway using systematically administrated barbiturates, which also have an agnostic effect on GABAergic receptors. Evoked potentials recorded from the cochlear nucleus and the inferior colliculus were substantially suppressed during pentobarbital anesthesia (40 mg/kg) when compared to awake recording (Webster & Aitkin, 1971). In auditory thalamus,

however, it was shown that pentobarbitone (35 – 40 mg/kg) caused less suppression when compared to auditory cortex (Makela, Karmos, Molnar, Csepe, & Winkler, 1990), or even facilitation (Webster & Aitkin, 1971). In rat auditory thalamus or primary auditory cortex, the suppressive effect of bolus pentobarbitone on neuronal activities were characterized in rats under ketamine (Wehr & Zador, 2005) or nitrous oxide anesthesia (Zurita, Villa, de Ribaupierre, de Ribaupierre, & Rouiller, 1994) in terms of response pattern and stimulus selectivity. Intracellular recording demonstrated a dramatic prolongation of sound-induced inhibitory conductance caused by bolus pentobarbitone (50 mg/kg) (Wehr & Zador, 2005), which suggested a large involvement of GABAergic inhibition in the auditory response generated in cortex.

GABAergic transmission participated in the inhibition of visual responses as shown in retinal rod bipolar cells (Eggers & Lukasiewicz, 2006), retinal ganglion neurons (Priest, Robbins, & Ikeda, 1985), and visual cortex (Kraut, Arezzo, & Vaughan, 1990). Although our results showed that middle-latency and long-latency VEPs were substantially attenuated during the 2%-isoflurane block, we did not achieve as much suppression in AEPs using isoflurane as expected from the barbiturate studies. Aside from the likely deeper anesthesia using pentobarbitone, the less suppression in AEPs in the current study is also possible as isoflurane may work through a non-GABAergic mechanism.

Due to the mechanism of GABAergic inhibition, it is speculated that multisensory interaction is also affected during anesthesia. Under propofol-sedation, there was a decrease in the functional integration between different neural networks revealed by resting-state fMRI (Schrouff et al., 2011), suggesting a suppressed cross-modal interaction. However, it has been recently shown that multisensory exposure under ketamine anesthesia could still improve the performance of visual localization in a cat model of hemianopia (Jiang, Rowland, & Stein, 2020; Yu, Rowland,

& Stein, 2010). Ketamine acts through NMDA receptors and therefore has a different mechanism from propofol/isoflurane and dexmedetomidine. Therefore, by using different anesthesia protocols, it is possible to contrast the involvement of different neuromodulation systems in cross-modal interactions.

6.5.3. Future studies

Using cat as an animal model, there are several directions where the findings of multiple temporal integrations in this thesis can be expanded in future studies. First of all, the effect of audiovisual temporal disparity can be examined in cats with hearing loss. In cats with partial hearing loss, the auditory cortex can still be activated with a click stimulus using an increased sound level. Low-frequency narrowband noise or tone pip can be considered as an alternative to click in the case of high-frequency hearing loss. Increased excitability has been shown in human auditory cortex after age-related hearing loss, suggesting a reduced cortical inhibition (Alain, Roye, & Salloum, 2014). Deficit in audiovisual integration was observed in subjects with hearing loss using auditory evoked potentials (Musacchia, Arum, Nicol, Garstecki, & Kraus, 2009). Studies on how partial hearing loss affects multisensory processing in the auditory cortex will help a better understanding of the speech communication challenge faced by the geriatric population.

In cats with profound hearing loss or deafness, cochlear implants can be performed to restore the inputs to the auditory nerves. Alternatively, electric stimulation of ventral part of medial geniculate body (MGBv) can evoke responses in the primary auditory cortex (A1). It has been shown that projections from the ventral division of the medial geniculate body (MGBv) into A1 were partially (about 40%) preserved in early-deaf cats (Chabot et al., 2015). Both techniques, in

the context of electric hearing, will enable us to investigate visual modulation in the auditory cortex. Making contrast between the two techniques can further differentiate audiovisual interaction between cortical and subcortical networks.

The onset and duration of deafness are known to affect the pattern of cross-modal plasticity, suggesting that the timing of brain development and auditory deprivation play an important role in the outcome of hearing restoration through cochlear implant (Ruff et al., 2017). Hearing loss or deafness can be induced in cats at any development stage (R. D. Brown & McElwee Jr, 1972; Sewell, 1984). As a follow-up study to this thesis, the audiovisual interaction can be quantified at any time point after deafening, using both extracellular recordings from chronically implanted electrodes and/or repeated non-invasive scalp-recorded auditory evoked potentials. Characterizing the temporal trajectory of the cross-modal plasticity and the re-establish of multisensory processing following deafness can be a critical step for optimizing the auditory training after cochlear implant.

6.6. References

Alain, C., Roye, A., & Salloum, C. (2014). Effects of age-related hearing loss and background noise on neuromagnetic activity from auditory cortex. *Frontiers in Systems Neuroscience, 8*, 8.

Anderson, C. A., Lazard, D. S., & Hartley, D. E. (2017). Plasticity in bilateral superior temporal cortex: Effects of deafness and cochlear implantation on auditory and visual speech processing. *Hearing Research, 343*, 138-149.

Anderson, C. A., Wiggins, I. M., Kitterick, P. T., & Hartley, D. E. (2017). Adaptive benefit of cross-modal plasticity following cochlear implantation in deaf adults. *Proceedings of the National Academy of Sciences, 114*(38), 10256-10261.

Atilgan, H., Town, S. M., Wood, K. C., Jones, G. P., Maddox, R. K., Lee, A. K., & Bizley, J. K. (2018). Integration of visual information in auditory cortex promotes auditory scene analysis through multisensory binding. *Neuron, 97*(3), 640-655. e644.

Barbour, D. L., & Wang, X. (2003). Contrast tuning in auditory cortex. *Science, 299*(5609), 1073-1075.

Barrow, P. C., & Schmitt, G. (2017). Juvenile nonclinical safety studies in support of pediatric drug development. *Drug Safety Evaluation: Methods and Protocols*, 25-67.

Bawazeer, N., Vuong, H., Riehm, S., Veillon, F., & Charpiot, A. (2019). Magnetic resonance imaging after cochlear implants. *Journal of Otology, 14*(1), 22-25.

Beauchamp, M. S., Nath, A. R., & Pasalar, S. (2010). fMRI-guided transcranial magnetic stimulation reveals that the superior temporal sulcus is a cortical locus of the McGurk effect. *Journal of Neuroscience, 30*(7), 2414-2417.

Berardi, N., Pizzorusso, T., & Maffei, L. (2000). Critical periods during sensory development. *Current Opinion in Neurobiology, 10*(1), 138-145.

Bertelson, P. (1967). The time course of preparation. *Quarterly Journal of Experimental Psychology, 19*(3), 272-279.

Bizley, J. K., Nodal, F. R., Bajo, V. M., Nelken, I., & King, A. J. (2007). Physiological and anatomical evidence for multisensory interactions in auditory cortex. *Cerebral Cortex, 17*(9), 2172-2189.

Bremen, P., Massoudi, R., Van Wanrooij, M. M., & Van Opstal, A. J. (2017). Audio-visual integration in a redundant target paradigm: a comparison between rhesus macaque and man. *Frontiers in Systems Neuroscience, 11*, 89.

Brown, R. D., & McElwee Jr, T. W. (1972). Effects of intra-arterially and intravenously administered ethacrynic acid and furosemide on cochlear N1 in cats. *Toxicology and Applied Pharmacology, 22*(4), 589-594.

Brown, T. A., Gati, J. S., Hughes, S. M., Nixon, P. L., Menon, R. S., & Lomber, S. G. (2014). Functional imaging of auditory cortex in adult cats using high-field fMRI. *JoVE (Journal of Visualized Experiments)*(84), e50872.

Brown, T. A., Joanisse, M. F., Gati, J. S., Hughes, S. M., Nixon, P. L., Menon, R. S., & Lomber, S. G. (2013). Characterization of the blood-oxygen level-dependent (BOLD) response in cat auditory cortex using high-field fMRI. *Neuroimage, 64*, 458-465.

Butler, B. E., Chabot, N., & Lomber, S. G. (2016). Quantifying and comparing the pattern of thalamic and cortical projections to the posterior auditory field in hearing and deaf cats. *Journal of Comparative Neurology, 524*(15), 3042-3063.

Butler, B. E., de la Rua, A., Ward-Able, T., & Lomber, S. G. (2018). Cortical and thalamic connectivity to the second auditory cortex of the cat is resilient to the onset of deafness. *Brain Structure and Function, 223*, 819-835.

Butler, B. E., Hall, A. J., & Lomber, S. G. (2015). High-field functional imaging of pitch processing in auditory cortex of the cat. *PloS One, 10*(7), e0134362.

Butler, B. E., Sunstrum, J. K., & Lomber, S. G. (2018). Modified origins of cortical projections to the superior colliculus in the deaf: Dispersion of auditory efferents. *Journal of Neuroscience, 38*(16), 4048-4058.

Camalier, C. R., D'Angelo, W. R., Sterbing-D'Angelo, S. J., de la Mothe, L. A., & Hackett, T. A. (2012). Neural latencies across auditory cortex of macaque support a dorsal stream supramodal timing advantage in primates. *Proceedings of the National Academy of Sciences, 109*(44), 18168-18173.

Caron-Desrochers, L., Schönwiesner, M., Focke, K., & Lehmann, A. (2018). Assessing visual modulation along the human subcortical auditory pathway. *Neuroscience Letters, 685*, 12-17.

Chabot, N., Butler, B. E., & Lomber, S. G. (2015). Differential modification of cortical and thalamic projections to cat primary auditory cortex following early-and late-onset deafness. *Journal of Comparative Neurology, 523*(15), 2297-2320.

Chen, L.-C., Sandmann, P., Thorne, J. D., Bleichner, M. G., & Debener, S. (2016). Cross-modal functional reorganization of visual and auditory cortex in adult cochlear implant users identified with fNIRS. *Neural Plasticity, 2016*.

Clemo, H. R., Lomber, S. G., & Meredith, M. A. (2016). Synaptic Basis for Cross-modal Plasticity: Enhanced Supragranular Dendritic Spine Density in Anterior Ectosylvian Auditory Cortex of the Early Deaf Cat. *Cerebral Cortex, 26*(4), 1365-1376. doi:10.1093/cercor/bhu225

Clemo, H. R., Lomber, S. G., & Meredith, M. A. (2017). Synaptic distribution and plasticity in primary auditory cortex (A1) exhibits laminar and cell-specific changes in the deaf. *Hearing Research, 353*, 122-134. doi:10.1016/j.heares.2017.06.009

Cozzi, B., Ballarin, C., Mantovani, R., & Rota, A. (2017). Aging and veterinary care of cats, dogs, and horses through the records of three university veterinary hospitals. *Frontiers in veterinary science, 4*, 14.

Crosse, M. J., Di Liberto, G. M., & Lalor, E. C. (2016). Eye can hear clearly now: inverse effectiveness in natural audiovisual speech processing relies on long-term crossmodal temporal integration. *Journal of Neuroscience, 36*(38), 9888-9895.

Dorrn, A. L., Yuan, K., Barker, A. J., Schreiner, C. E., & Froemke, R. C. (2010). Developmental sensory experience balances cortical excitation and inhibition. *Nature, 465*(7300), 932-936.

Eggermont, J. J. (1993). Differential effects of age on click-rate and amplitude modulation-frequency coding in primary auditory cortex of the cat. *Hearing Research, 65*(1-2), 175-192.

Eggers, E. D., & Lukasiewicz, P. D. (2006). GABA(A), GABA(C) and glycine receptor-mediated inhibition differentially affects light-evoked signalling from mouse retinal rod bipolar cells. *J Physiol, 572*(Pt 1), 215-225. doi:10.1113/jphysiol.2005.103648

Fernández-Ruiz, A., Muñoz, S., Sancho, M., Makarova, J., Makarov, V. A., & Herreras, O. (2013). Cytoarchitectonic and dynamic origins of giant positive local field potentials in the dentate gyrus. *Journal of Neuroscience, 33*(39), 15518-15532.

Fort, A., Delpuech, C., Pernier, J., & Giard, M.-H. (2002). Dynamics of cortico-subcortical cross-modal operations involved in audio-visual object detection in humans. *Cerebral Cortex, 12*(10), 1031-1039.

Foss-Feig, J. H., Kwakye, L. D., Cascio, C. J., Burnette, C. P., Kadivar, H., Stone, W. L., & Wallace, M. T. (2010). An extended multisensory temporal binding window in autism spectrum disorders. *Experimental Brain Research, 203*, 381-389.

Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. y. (2004). Recalibration of audiovisual simultaneity. *Nature Neuroscience, 7*(7), 773-778.

Ghazanfar, A. A., Maier, J. X., Hoffman, K. L., & Logothetis, N. K. (2005). Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *Journal of Neuroscience, 25*(20), 5004-5012.

Giard, M. H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *Journal of Cognitive Neuroscience, 11*(5), 473-490.

Gourévitch, B., & Eggermont, J. J. (2008). Spectro-temporal sound density-dependent long-term adaptation in cat primary auditory cortex. *European Journal of Neuroscience, 27*(12), 3310-3321.

Hackett, T. A. (2011). Information flow in the auditory cortical network. *Hearing Research, 271*(1-2), 133-146.

Hairston, W. D., Burdette, J. H., Flowers, D. L., Wood, F. B., & Wallace, M. T. (2005). Altered temporal profile of visual–auditory multisensory interactions in dyslexia. *Experimental Brain Research, 166*, 474-480.

Hall, A. J., Brown, T. A., Grahn, J. A., Gati, J. S., Nixon, P. L., Hughes, S. M., . . . Lomber, S. G. (2014). There's more than one way to scan a cat: imaging cat auditory cortex with high-field fMRI using continuous or sparse sampling. *Journal of Neuroscience Methods, 224*, 96-106.

Hall, A. J., Butler, B. E., & Lomber, S. G. (2016). The cat's meow: A high-field fMRI assessment of cortical activity in response to vocalizations and complex auditory stimuli. *Neuroimage, 127*, 44-57.

Hall, A. J., & Lomber, S. G. (2015). High-field fMRI reveals tonotopically-organized and core auditory cortex in the cat. *Hearing Research, 325*, 1-11.

Harrar, V., Harris, L. R., & Spence, C. (2017). Multisensory integration is independent of perceived simultaneity. *Experimental Brain Research, 235*, 763-775.

Harrison, N. L., Kugler, J. L., Jones, M. V., Greenblatt, E. P., & Pritchett, D. B. (1993). Positive modulation of human gamma-aminobutyric acid type A and glycine receptors by the inhalation anesthetic isoflurane. *Mol Pharmacol, 44*(3), 628-632.

Herreras, O. (2016). Local field potentials: myths and misunderstandings. *Frontiers in neural circuits, 10*, 101.

Hillock-Dunn, A., & Wallace, M. T. (2012). Developmental changes in the multisensory temporal binding window persist into adolescence. *Developmental science, 15*(5), 688-696.

HOYUMPA, A., RODAN, I., Brown, M., BROWN, S., BUFFINGTON, T., LARUE, M., . . . SPARKES, A. (2010). Feline Life Stage Guidelines. *Journal of the American Animal Hospital Association, 46*(1), 70-71.

Jia, F., Yue, M., Chandra, D., Homanics, G. E., Goldstein, P. A., & Harrison, N. L. (2008). Isoflurane is a potent modulator of extrasynaptic GABAA receptors in the thalamus. *J Pharmacol Exp Ther, 324*(3), 1127-1135. doi:10.1124/jpet.107.134569

Jiang, H., Rowland, B. A., & Stein, B. E. (2020). Reversing hemianopia by multisensory training under anesthesia. *Frontiers in Systems Neuroscience*, 4.

Kajikawa, Y., & Schroeder, C. E. (2011). How local is the local field potential? *Neuron, 72*(5), 847-858.

Kawase, T., Yahata, I., Kanno, A., Sakamoto, S., Takanashi, Y., Takata, S., . . . Katori, Y. (2016). Impact of audio-visual asynchrony on lip-reading effects-neuromagnetic and psychophysical study. *PloS One, 11*(12), e0168740.

Kayser, C., Petkov, C. I., & Logothetis, N. K. (2008). Visual modulation of neurons in auditory cortex. *Cerebral Cortex, 18*(7), 1560-1574.

Khandhadia, A. P., Murphy, A. P., Romanski, L. M., Bizley, J. K., & Leopold, D. A. (2021). Audiovisual integration in macaque face patch neurons. *Current Biology, 31*(9), 1826-1835. e1823.

Kok, M. A., Chabot, N., & Lomber, S. G. (2014). Cross-modal reorganization of cortical afferents to dorsal auditory cortex following early-and late-onset deafness. *Journal of Comparative Neurology, 522*(3), 654-675.

Krasowski, M. D., Koltchine, V. V., Rick, C. E., Ye, Q., Finn, S. E., & Harrison, N. L. (1998). Propofol and other intravenous anesthetics have sites of action on the γ-aminobutyric acid type A receptor distinct from that for isoflurane. *Mol Pharmacol, 53*(3), 530-538. doi:10.1124/mol.53.3.530

Kraut, M. A., Arezzo, J. C., & Vaughan, H. G., Jr. (1990). Inhibitory processes in the flash evoked potential of the monkey. *Electroencephalogr Clin Neurophysiol, 76*(5), 440-452. doi:10.1016/0013-4694(90)90097-4

Lakatos, P., Chen, C.-M., O'Connell, M. N., Mills, A., & Schroeder, C. E. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron, 53*(2), 279-292.

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science, 320*(5872), 110-113.

Lebib, R., Papo, D., de Bode, S., & Baudonnière, P.-M. (2003). Evidence of a visual-to-auditory cross-modal sensory gating phenomenon as reflected by the human P50 event-related brain potential modulation. *Neuroscience Letters, 341*(3), 185-188.

Lee, C. C., & Winer, J. A. (2008). Connections of cat auditory cortex: III. Corticocortical system. *Journal of Comparative Neurology, 507*(6), 1920-1943.

Levine, A. T., Li, B., Barnes, P., Lomber, S. G., & Butler, B. E. (2020). Assessment of anesthesia on physiological stability and BOLD signal reliability during visual or acoustic stimulation in the cat. *Journal of Neuroscience Methods, 334*, 108603.

Li, B., Li, B. W., Chen, Y., Wang, L. H., & Diao, Y. C. (2000). Response properties of PMLS and PLLS neurons to simulated optic flow patterns. *European Journal of Neuroscience, 12*(5), 1534-1544.

Lomber, S. G. (1999). The advantages and limitations of permanent or reversible deactivation techniques in the assessment of neural function. *Journal of Neuroscience Methods, 86*(2), 109-117.

Lomber, S. G., Meredith, M. A., & Kral, A. (2010). Cross-modal plasticity in specific auditory cortices underlies visual compensations in the deaf. *Nature Neuroscience, 13*(11), 1421-1427.

Lomber, S. G., Payne, B. R., & Horel, J. A. (1999). The cryoloop: an adaptable reversible cooling deactivation method for behavioral or electrophysiological assessment of neural function. *Journal of Neuroscience Methods, 86*(2), 179-194.

Los, S. A., & Van Den Heuvel, C. E. (2001). Intentional and unintentional contributions to nonspecific preparation during reaction time foreperiods. *Journal of Experimental Psychology: Human Perception and Performance, 27*(2), 370.

Makela, J. P., Karmos, G., Molnar, M., Csepe, V., & Winkler, I. (1990). Steady-state responses from the cat auditory cortex. *Hear Res, 45*(1-2), 41-50. doi:10.1016/0378-5955(90)90181-n

McDonald, J. J., Teder-SaÈlejaÈrvi, W. A., & Hillyard, S. A. (2000). Involuntary orienting to sound improves visual perception. *Nature, 407*(6806), 906-908.

McGovern, D. P., Roudaia, E., Newell, F. N., & Roach, N. W. (2016). Perceptual learning shapes multisensory causal inference via two distinct mechanisms. *Scientific Reports, 6*(1), 24673.

Mégevand, P., Molholm, S., Nayak, A., & Foxe, J. J. (2013). Recalibration of the multisensory temporal window of integration results from changing task demands. *PloS One, 8*(8), e71608.

Meredith, M. A., & Allman, B. L. (2015). Single-unit analysis of somatosensory processing in the core auditory cortex of hearing ferrets. *European Journal of Neuroscience, 41*(5), 686-698. doi:10.1111/ejn.12828

Meredith, M. A., & Allman, B. L. (2015). Single-unit analysis of somatosensory processing in the core auditory cortex of hearing ferrets. *European Journal of Neuroscience, 41*(5), 686-698.

Meredith, M. A., Kryklywy, J., McMillan, A. J., Malhotra, S., Lum-Tai, R., & Lomber, S. G. (2011). Crossmodal reorganization in the early deaf switches sensory, but not behavioral roles of auditory cortex. *Proceedings of the National Academy of Sciences, 108*(21), 8856-8861.

Meredith, M. A., & Lomber, S. G. (2011). Somatosensory and visual crossmodal plasticity in the anterior auditory field of early-deaf cats. *Hearing Research, 280*(1-2), 38-47.

Meredith, M. A., Nemitz, J. W., & Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *Journal of Neuroscience, 7*(10), 3215-3229.

Merrikhi, Y., Kok, M. A., Carrasco, A., Meredith, M. A., & Lomber, S. G. (2022). Multisensory responses in a belt region of the dorsal auditory cortical pathway. *European Journal of Neuroscience, 55*(2), 589-610.

Merrikhi, Y., Kok, M. A., Lomber, S. G., & Meredith, M. A. (2023). A comparison of multisensory features of two auditory cortical areas: primary (A1) and higher-order dorsal zone (DZ). *Cerebral Cortex Communications, 4*(1), tgac049.

Miller, J. (1986). Timecourse of coactivation in bimodal divided attention. *Perception and Psychophysics, 40*(5), 331-343.

Miller, L. M., Escabı, M. A., Read, H. L., & Schreiner, C. E. (2001). Functional convergence of response properties in the auditory thalamocortical system. *Neuron, 32*(1), 151-160.

Musacchia, G., Arum, L., Nicol, T., Garstecki, D., & Kraus, N. (2009). Audiovisual deficits in older adults with hearing loss: biological evidence. *Ear and Hearing, 30*(5), 505-514.

Musacchia, G., Sams, M., Nicol, T., & Kraus, N. (2006). Seeing speech affects acoustic information processing in the human brainstem. *Experimental Brain Research, 168*, 1-10.

Nakahiro, M., Yeh, J. Z., Brunner, E., & Narahashi, T. (1989). General anesthetics modulate GABA receptor channel complex in rat dorsal root ganglion neurons. *FASEB J, 3*(7), 1850-1854. doi:10.1096/fasebj.3.7.2541038

Nardone, R., Florea, C., Höller, Y., Brigo, F., Versace, V., Lochner, P., . . . Trinka, E. (2017). Rodent, large animal and non-human primate models of spinal cord injury. *Zoology, 123*, 101-114.

Navarra, J., Fernández-Prieto, I., & Garcia-Morera, J. (2013). Realigning thunder and lightning: temporal adaptation to spatiotemporally distant events. *PloS One, 8*(12), e84278.

Navarra, J., Vatakis, A., Zampini, M., Soto-Faraco, S., Humphreys, W., & Spence, C. (2005). Exposure to asynchronous audiovisual speech extends the temporal window for audiovisual integration. *Cognitive Brain Research, 25*(2), 499-507.

Niemi, P., & Näätänen, R. (1981). Foreperiod and simple reaction time. *Psychological Bulletin, 89*(1), 133.

Noel, J.-P., De Niear, M., Van der Burg, E., & Wallace, M. T. (2016). Audiovisual simultaneity judgment and rapid recalibration throughout the lifespan. *PloS One, 11*(8), e0161698.

Nozaradan, S., Peretz, I., & Mouraux, A. (2012). Steady-state evoked potentials as an index of multisensory temporal binding. *Neuroimage, 60*(1), 21-28.

Olejniczak, P. (2006). Neurophysiologic basis of EEG. *Journal of Clinical Neurophysiology, 23*(3), 186-189.

Oray, S., Lu, Z.-L., & Dawson, M. E. (2002). Modification of sudden onset auditory ERP by involuntary attention to visual stimuli. *International Journal of Psychophysiology, 43*(3), 213-224.

Palmer, L., Rosenquist, A., & Tusa, R. (1978). The retinotopic organization of lateral suprasylvian visual areas in the cat. *Journal of Comparative Neurology, 177*(2), 237-256.

Parise, C. V., Harrar, V., Ernst, M. O., & Spence, C. (2013). Cross-correlation between auditory and visual signals promotes multisensory integration. *Multisensory research, 26*(3), 307-316.

Parise, C. V., Spence, C., & Ernst, M. O. (2012). When correlation implies causation in multisensory integration. *Current Biology, 22*(1), 46-49.

Perrodin, C., Kayser, C., Logothetis, N. K., & Petkov, C. I. (2015). Natural asynchronies in audiovisual communication signals regulate neuronal multisensory interactions in voice-sensitive cortex. *Proceedings of the National Academy of Sciences, 112*(1), 273-278.

Phillips, D., & Orman, S. (1984). Responses of single neurons in posterior field of cat auditory cortex to tonal stimulation. *Journal of Neurophysiology, 51*(1), 147-163.

Phillips, D., Orman, S., Musicant, A., & Wilson, G. (1985). Neurons in the cat's primary auditory cortex distinguished by their responses to tones and wide-spectrum noise. *Hearing Research, 18*(1), 73-86.

Posner, M. I., & Boies, S. J. (1971). Components of attention. *Psychological Review, 78*(5), 391.

Priest, T. D., Robbins, J., & Ikeda, H. (1985). The action of inhibitory neurotransmitters, γ-aminobutyric acid and glycine may distinguish between the area centralis and the peripheral retina in cats. *Vision Res, 25*(12), 1761-1770. doi:10.1016/0042-6989(85)90001-x

Rabinowitz, N. C., Willmore, B. D. B., Schnupp, J. W. H., & King, A. J. (2011). Contrast Gain Control in Auditory Cortex. *Neuron, 70*(6), 1178-1191. doi:10.1016/j.neuron.2011.04.030

Rubenstein, J., & Merzenich, M. M. (2003). Model of autism: increased ratio of excitation/inhibition in key neural systems. *Genes, Brain and Behavior, 2*(5), 255-267.

Ruff, S., Bocklet, T., Nöth, E., Müller, J., Hoster, E., & Schuster, M. (2017). Speech production quality of cochlear implant users with respect to duration and onset of hearing loss. *ORL, 79*(5), 282-294.

Schormans, A. L., Scott, K. E., Vo, A. M., Tyker, A., Typlt, M., Stolzberg, D., & Allman, B. L. (2017). Audiovisual temporal processing and synchrony perception in the rat. *Frontiers in Behavioral Neuroscience, 10*, 246.

Schreiner, C. E., & Winer, J. A. (2007). Auditory cortex mapmaking: principles, projections, and plasticity. *Neuron, 56*(2), 356-365.

Schrouff, J., Perlbarg, V., Boly, M., Marrelec, G., Boveroux, P., Vanhaudenhuyse, A., . . . Pélégrini-Issac, M. (2011). Brain functional integration decreases during propofol-induced loss of consciousness. *Neuroimage, 57*(1), 198-205.

Sewell, W. F. (1984). The effects of furosemide on the endocochlear potential and auditory-nerve fiber tuning curves in cats. *Hearing Research, 14*(3), 305-314.

Stekelenburg, J. J., & Vroomen, J. (2007). Neural correlates of multisensory integration of ecologically valid audiovisual events. *Journal of Cognitive Neuroscience, 19*(12), 1964-1973.

Stevenson, R. A., & Wallace, M. T. (2013). Multisensory temporal integration: task and stimulus dependencies. *Experimental Brain Research, 227*, 249-261.

Stevenson, R. A., Zemtsov, R. K., & Wallace, M. T. (2012). Individual differences in the multisensory temporal binding window predict susceptibility to audiovisual illusions. *Journal of Experimental Psychology: Human Perception and Performance, 38*(6), 1517.

Stolzberg, D., Butler, B. E., & Lomber, S. G. (2018). Effects of neonatal deafness on resting-state functional network connectivity. *Neuroimage, 165*, 69-82.

Stolzberg, D., Wong, C., Butler, B. E., & Lomber, S. G. (2017). Catlas: an magnetic resonance imaging-based three-dimensional cortical atlas and tissue probability maps for the domestic cat (Felis catus). *Journal of Comparative Neurology, 525*(15), 3190-3206.

Sutter, M., Schreiner, C., McLean, M., O'connor, K., & Loftus, W. (1999). Organization of inhibitory frequency receptive fields in cat primary auditory cortex. *Journal of Neurophysiology, 82*(5), 2358-2371.

Sutter, M. L., & Schreiner, C. E. (1991). Physiology and topography of neurons with multipeaked tuning curves in cat primary auditory cortex. *Journal of Neurophysiology, 65*(5), 1207-1226.

Tajiri, N., Dailey, T., Metcalf, C., Mosley, Y. I., Lau, T., Staples, M., . . . Yasuhara, T. (2013). In vivo animal stroke models: a rationale for rodent and non-human primate models. *Translational stroke research, 4*, 308-321.

Takei, T., Lomber, S. G., Cook, D. J., & Scott, S. H. (2021). Transient deactivation of dorsal premotor cortex or parietal area 5 impairs feedback control of the limb in macaques. *Current Biology, 31*(7), 1476-1487. e1475.

Uğurbil, K. (2021). Ultrahigh field and ultrahigh resolution fMRI. *Current opinion in biomedical engineering, 18*, 100288.

Valentine, P. A., & Eggermont, J. J. (2004). Stimulus dependence of spectro-temporal receptive fields in cat primary auditory cortex. *Hearing Research, 196*(1-2), 119-133.

Van der Burg, E., & Goodbourn, P. T. (2015). Rapid, generalized adaptation to asynchronous audiovisual speech. *Proceedings of the Royal Society B: Biological Sciences, 282*(1804), 20143083.

Van der Burg, E., Olivers, C. N., Bronkhorst, A. W., & Theeuwes, J. (2008). Pip and pop: nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology: Human Perception and Performance, 34*(5), 1053.

Vatakis, A., Navarra, J., Soto-Faraco, S., & Spence, C. (2007). Temporal recalibration during asynchronous audiovisual speech perception. *Experimental Brain Research, 181*, 173-181.

Wang, X. (2007). Neural coding strategies in auditory cortex. *Hearing Research, 229*(1-2), 81-93.

Webster, W. R., & Aitkin, L. M. (1971). Evoked potential and single unit studies of neural mechanisms underlying the effects of repetitive stimulation in the auditory pathway. *Electroencephalogr Clin Neurophysiol, 31*(6), 581-592. doi:10.1016/0013-4694(71)90074-5

Wehr, M., & Zador, A. M. (2005). Synaptic mechanisms of forward suppression in rat auditory cortex. *Neuron, 47*(3), 437-445. doi:10.1016/j.neuron.2005.06.009

Winer, J. A., Miller, L. M., Lee, C. C., & Schreiner, C. E. (2005). Auditory thalamocortical transformation: structure and function. *Trends in Neurosciences, 28*(5), 255-263.

Wong, C., Chabot, N., Kok, M. A., & Lomber, S. G. (2015). Amplified somatosensory and visual cortical projections to a core auditory area, the anterior auditory field, following early-and late-onset deafness. *Journal of Comparative Neurology, 523*(13), 1925-1947.

Wong, C., Pearson, K. G., & Lomber, S. G. (2018). Contributions of parietal cortex to the working memory of an obstacle acquired visually or tactilely in the locomoting cat. *Cerebral Cortex, 28*(9), 3143-3158.

Yang, J., Isenberg, K. E., & Zorumski, C. F. (1992). Volatile anesthetics gate a chloride current in postnatal rat hippocampal neurons. *FASEB J, 6*(3), 914-918. doi:10.1096/fasebj.6.3.1740240

Yu, L., Rowland, B. A., & Stein, B. E. (2010). Initiating the development of multisensory integration by manipulating sensory experience. *Journal of Neuroscience, 30*(14), 4904-4913.

Yuan, Y., Lleo, Y., Daniel, R., White, A., & Oh, Y. (2021). The impact of temporally coherent visual cues on speech perception in Complex auditory environments. *Frontiers in Neuroscience, 15*, 678029.

Yuan, Y., Wayland, R., & Oh, Y. (2020). Visual analog of the acoustic amplitude envelope benefits speech perception in noise. *Journal of the Acoustical Society of America, 147*(3), El246. doi:10.1121/10.0000737

Zhang, Y., Stabernack, C., Sonner, J., Dutton, R., & Eger, E. I. (2001). Both cerebral GABAA receptors and spinal GABAA receptors modulate the capacity of isoflurane to produce immobility. *Anesthesia & Analgesia, 92*(6), 1585-1589. doi:10.1097/00000539-200106000-00047

Zheng, W., & Knudsen, E. I. (1999). Functional selection of adaptive auditory space map by GABAA-mediated inhibition. *Science, 284*(5416), 962-965.

Zurita, P., Villa, A. E., de Ribaupierre, Y., de Ribaupierre, F., & Rouiller, E. M. (1994). Changes of single unit activity in the cat's auditory thalamus and cortex associated to different anesthetic conditions. *Neurosci Res, 19*(3), 303-316. doi:10.1016/0168-0102(94)90043-4