“Laurel & Yanny”:
EEG Neural Correlates of an Auditory Bistable Language Stimulus

A Thesis
Presented to
The Established Interdisciplinary Committee for Neuroscience
Reed College

In Partial Fulfillment
of the Requirements for the Degree
Bachelor of Arts

Alex Franklin
December 2021

Approved for the Committee

(Neuroscience)

Michael Pitts

Acknowledgments

I've taken a personal moment below to thank friends and family who have supported and helped me in life, college, and through the thesis process.

Nico, you are my brother and my absolute best friend. Thank you for being my sibling in this life. Your support means the world to me. I am always here for you, and I love you so much.

Mama, te quiero mucho. Gracias por tu sabiduría y tus abrazos. Tu humor y tu amor incondicional han sido los mayores regalos en mi vida.

Dad, thank you for the inspiration of a lifetime. The sound of your music from mom's arms are some of my earliest memories. Your sweaters and socks are mine. I love you.

Maddi. You are a friend of the world. I could never begin to thank you enough for your love, intuition, wisdom, support, and beautiful music. You are my person. I love you and Moglie always.

Micah, Amaury, Sam, Flynn, Kian, EJ, Maya, Eli and Emilio. Thank you. You are all so beautiful. Your friendship makes me cry, you are blessings on Earth. I love you so much.

David Egert, Jim, Vicky, Tío Gero, Cathy, and David Gaylord (as well as Michael, David Presti and all of the Newt Riders). Thank you so dearly for your love and guidance.

To Grandpa & Grandma – thank you. I love you and miss you always.

Farhan, Noah, Maddy, Declan, Aiko, Leila, Makana, Frank, Alex, Eliza, Lucas and Finn. Thank you for being my friends here at Reed. The perspectives, support, friendship, and absurdity you have gifted me are so very close to my heart.

Michael Pitts, thank you so, so much for your wonderful guidance, your compelling lessons, and for giving me the opportunity to do this project.

Kevin Ortego, your generous help on this thesis was a true blessing. I could not have done this project nearly as well without your wisdom; thank you so much.

Thank you so very much for your lessons and assistance Enriqueta, and thank you to the data@reed team for your help coding on R and to everyone from the lab who helped with capping participants.

Thank you to all who participated in this thesis study.

Table of Contents

List of Figures

Abstract

“Laurel & Yanny” is a short auditory clip that became internet famous in May of 2018 due its bewildering perceptual bistability. Bistable stimuli contain two perceptually unique stable states of interpretation. Perceptual systems cannot settle on one state, instead adopting one interpretation for an amount of time (“Laurel”) and inevitably undergoing a perceptual reversal to the other (“Yanny”) after repeated presentations of the physically unchanging stimulus; making bistable stimuli fitting research tools by which to study changes in conscious perception. Preliminary EEG studies of auditory bistability are limited compared to those studying visual bistability. The present study used an electroencephalography (EEG) paradigm to investigate the neural correlates of the auditory bistable language stimulus “Laurel & Yanny”.

Results revealed three event-related potentials (ERPs) linked with perceptual reversals: a Mismatch-Negativity like component (MMN), followed by an auditory Reversal Negativity (aRN) and an auditory Late Positive Complex (aLPC). These results replicate and support recent evidence suggesting that the aRN and aLPC are auditory analogs of the widely studied visual reversal negativity (RN) and Late Positive Complex (LPC) (Davidson & Pitts, 2014). The MMN-like effect, however, was an unexpected finding; as it has been previously claimed that the MMN cannot be elicited by a single stimulus and can only be elicited by an exogenously deviant stimulus within a repeated pattern of standard stimuli (Bartha-Doering et al., 2015). Results did not identify a neural difference between participants' reported perception of Laurel and Yanny, suggesting that a more sensitive form of neural recordings or analyses must be used to detect a difference between them; such as multivariate pattern classification analysis (“Decoding”) or intracranial EEG. The RN has been postulated to index perceptual changes closely linked with conscious experience of a stimulus. Some have argued that the LPC indexes processing supporting perceptual awareness, while others have proposed a closer link with post-perceptual processing, or the reporting of one’s perception. Attempts to interpret the neural mechanisms indexed by the aRN and aLPC are outlined in the discussion section.

Chapter 1: Introduction

1.1 Ambiguity? Are You Sure?

Our minds and bodies inhabit a mysterious world filled with an inconceivable magnitude of elaborate stimuli. In our daily lives, we presume to reliably perceive what we experience from the external world through our senses, which help us cultivate an internal narrative about our lives and perceptions. Although we may consider our internal perceptual narrative to be in physical alignment with the external world, at times we encounter scenarios or stimuli that seem to cheat our perception and mislead us to perceive things as they are not. Perceptual deceptions such as these can most commonly be referred to as and demonstrated by phenomena known as "illusions". Ordinarily referred to as "optical illusions", images such as the one below appear to have qualities that perceptually defy laws of physics.

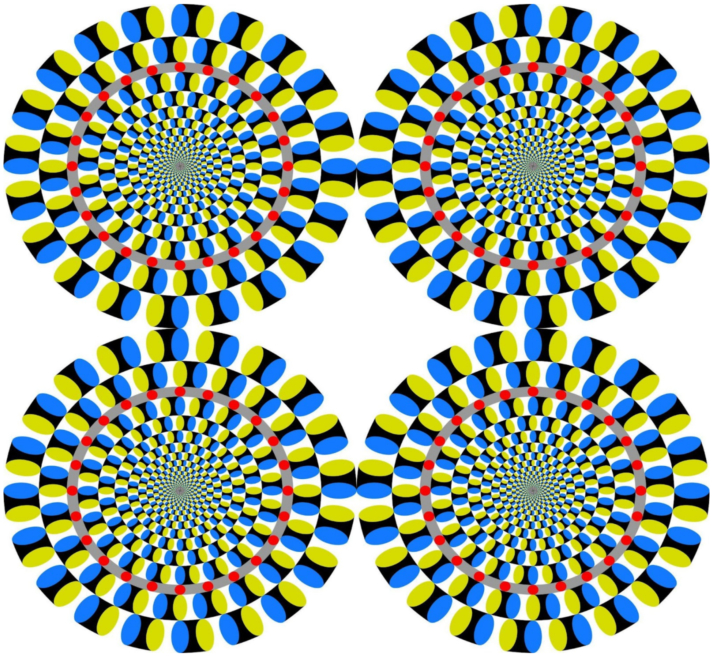


Figure 1.1 "Rotating-Snake" Visual Illusion.
The upper left and lower right circular patterns appear to move in counter-clockwise rotation, while the upper right and lower left appear to move in clockwise rotation (Source: Kitaoka, 2018).

Figure 1.1 depicts a version of the popular "rotating-snake illusion" (Kitaoka, 2018). When observers freely scan this illusion by making small and frequent eye movements, the upper left and lower right circular patterns appear to move in counterclockwise rotation, while the upper right and lower left appear to move in clockwise rotation (Kitaoka, 2018). In reality and in the most literal sense; the image is not moving, it's a static graphic pattern. The four circles are perfectly fixed on the page, however our sensory systems produce a perception of motion in the image (note that if you perfectly fixate your eyes in the center without making even tiny eye movements, perceived motion stops).

Optical illusions give us a uniquely profound opportunity to appreciate that our perception does not always coincide with the real world. However, this privilege is not always a guaranteed opportunity when we experience the world around us. Some scientists theorize that illusions give us a window in the status quo of perception, that is; almost everything we experience through our senses is somewhat of an illusion – "we don't see the physical world" (Howe & Purves, 2010). Many illusions may go unnoticed in the natural world, and indeed in our own narratives. How might we truly know? How could we even begin to attempt to understand such phenomena about our minds and our world? These are very baffling questions, ones that have inspired and mystified the minds of daydreaming children, neuroscientists, and philosophers for centuries.

Attempting to answer these challenging questions brings us to the topic of perceptual ambiguity, and its implications for the neuroscientific research of perception. Our opinions, emotions, memories and even our concrete witnessing of shared experiences often differ from those around us. This begs the question; how best can we study a fundamental and mutually objective internal experience; an experience that can reliably be perceived by each of our minds in the same way?

Perceptual ambiguity is a fitting candidate. Of course, we must define what is meant by "perceptual ambiguity"; for it could be difficult to imagine how perceiving ambiguity can even be a reliably shared experience as the magnitude of stimuli in our world may be indefinitely and conceptually ambiguous in nature. However, certain ambiguous stimuli solely influence one of our senses, whether visual or auditory (etc.) and have a defined number of stable states at which they are commonly experienced by us as individuals. Importantly, in any moment in time, when faced with ambiguous sensory input; our perceptual systems settle on one of (potentially many) "best guesses" presumably to help guide our actions and interactions with objects in the world. There exists a certain type of ambiguous stimuli that are known to be “bi-stable”. Commonly referred to as “Bistable Stimuli”, these stimuli inherently contain two perceptually unique stable interpretations. Our perceptual systems do not settle on solely perceiving one form of the stimulus, instead they adopt one interpretation for a certain amount of time (usually several seconds in duration) and then flip to the other perceptual interpretation; and then back to the first in a never-ending cycle despite the physical stimulus remaining unchanged. Bistable stimuli are thus valuable research tools for studying questions regarding our perception, because sensory input is completely held constant while we experience two (often very different) alternating perceptual interpretations. We can use bistable stimuli to study conscious perception and related psychological functions such as attention and memory, and when combined with scientific measurements of the brain; we can begin to understand how our minds interact with stimuli in our world.

There are at least two important comparisons that can be made when assessing brain activity underlying bistable perception. We can compare neural activity patterns when someone perceives one interpretation versus the other, and when one’s perception of the stimulus changes versus stays the same. Comparing neural activity patterns when someone perceives one interpretation versus the other can tell us something about the differential brain activity supporting the content of our conscious perception (while sensory input is held constant). Comparing when one’s perception of the stimulus changes versus stays the same can help reveal processes related to updating our conscious experience over time.

Experiencing a perceptual switch (often called a “reversal”) when attending to a bistable stimulus differs drastically from simply introducing one novel stimuli after another, which would instead document neural correlates indicative of changes in our sensory input. Although these two instances may have qualities in common with one another (both involving, influencing, and changing the contents of our attention and perception), experiencing a reversal while attending to a bistable stimulus gives us the opportunity to isolate brain activity affiliated with perceptual awareness.

1.2 Bistable Perception

1.2.1 Examples of Visual Bistability

Vision is the dominant sensory modality in humans and also the most widely studied in perception research. Visual figures are thus one of the clearest ways to exemplify perceptual bistability in a stimulus, as our eyes can remain fixated on one image; holding constant our sensory input while allowing our perceptual interpretations to alter under the guise of one still image.

Many stimuli that perfectly exemplify visual bistability are some of the most infamous optical illusions, such as the Rubin face-vase (Vercillo, 2012), Schroeder's stairs (Kitaoka, 2007), the Necker Cube (Maksim & Boivie, 2008) and the Jastow Rabbit-Duck (Ramachandran & Hirstein, 1997) depicted in Figure 1.2. These four images each have two mutually exclusive perceptual interpretations, ultimately appearing to be bistable with two commonly perceivable states of stability. One can examine each of these pictorial representations of bistability and attempt to witness both stable states as well as perceptual "reversals" between states.

These images have the astounding ability to alter our internal perceptual narratives. You may find that your perception of them inevitably switches on its own over time, with no conscious or volitional attempt to alter your own perception of them. You may find yourself having trouble viewing the alternative interpretation, and then suddenly experience an unexpected perceptual reversal. In another instance, you might successfully be able to fixate on one stable interpretation of the image, or rather be unable to switch at all! Observers also have a degree of voluntary control over what is perceived in these images (van Ee et al, 2005). For example, you can try to increase the rate at which you see each alternative, or try to hold one interpretation stable for as long as possible. While perhaps not completely under your control, you may often be able to influence what you see via volitional effort and attention, even when keeping your eyes fixated on the center of each image.

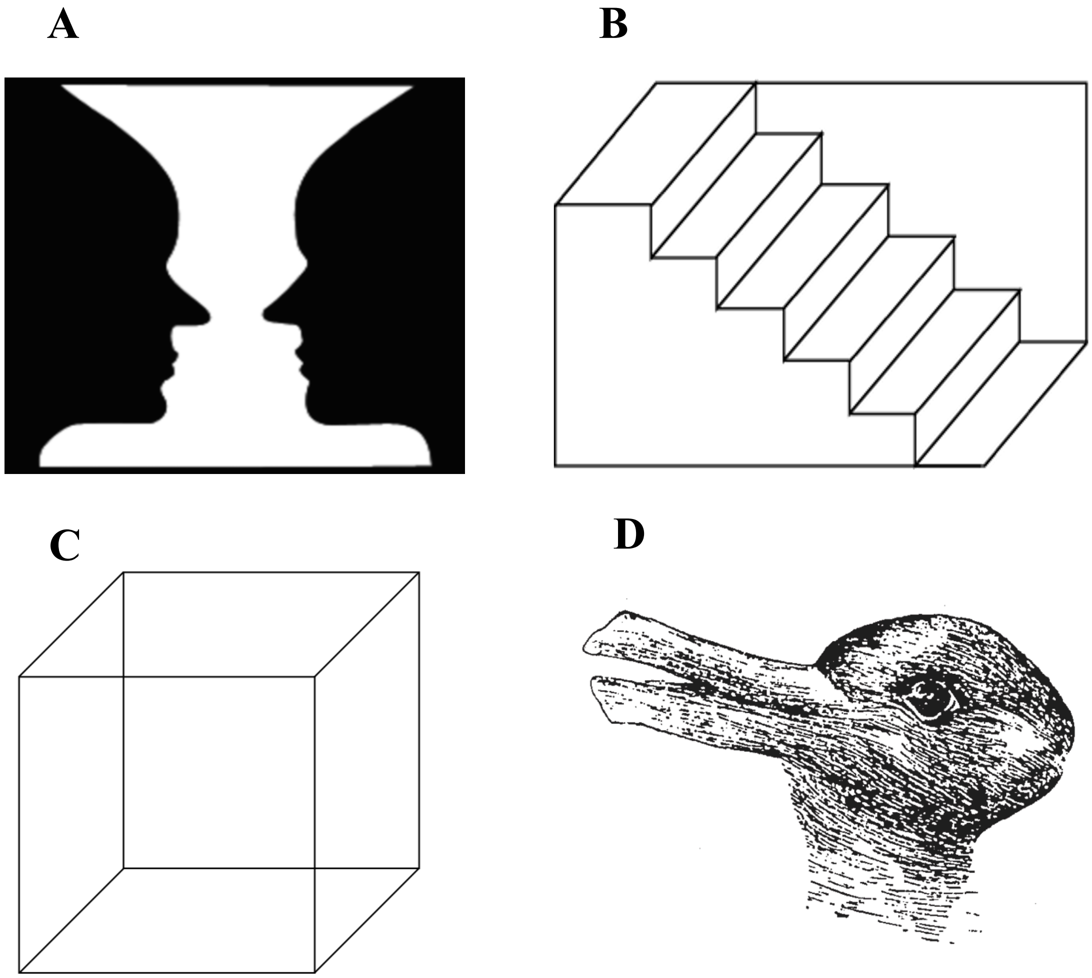


Figure 1.2 Examples of Visual Bistability.

(A) The Rubin face-vase. Created in 1915 by psychologist Edgar Rubin, the image displays a white vase positioned before a black background, also perceivable as two black silhouetted faces facing one another before a white background (Source: Vercillo, 2012). **(B)** Schroeder's staircase. The image displays a right-side up staircase, with the topmost step at the left and the bottom most step at the right. This image can also be perceived as an upside-down staircase, with the bottom-most step now positioned at the top left and the top-most step positioned at the bottom right (Source: Kitaoka, 2007). **(C)** The Necker Cube. The image presents a cube perceivable as either having the bottom left square-face in the front (with the upper-left face in the background; Source: Stevo-88, 2006) or as having the upper right square-face in the front (with the bottom-right face in the background). **(D)** The Jastow rabbit-duck, drawn by Joseph Jastow, who earned the first Ph.D. in Psychology ever awarded in the United States (El-Hai, n.d.). This image presents a hand-drawn image of the faces of both animals merged into one image (Source: Ramachandran & Hirstein, 1997).

Although many who look upon illusions or images such as these may regard them to "trick us" into perceiving something as if it is a lie, in actuality; our minds are not being cheated; nor are our brains "messing this one up". The mental processes underlying these instances of perceptual reversals abide by the same laws of perception that we experience in other real world environments and conditions. In order to answer complex personal questions and theories about the mechanisms underlying our attentional awareness, consciousness, and perception of the world; we can begin by examining such "low hanging fruit" so to speak: neural processes that take place when one experiences such perceptual phenomena as bistability.

1.2.2 Auditory Bistability and "Laurel & Yanny"

Some of the oldest auditory illusions are theorized by researchers to be echoes and thunderous reverberations, thought to be documented in ancient prehistoric art depicting stomping hooved animals painted upon the wall of a highly reverbic cave in India (Acoustical Society of America, 2014). Auditory illusions are some of the most extraordinary forms of illusory stimuli: A name one just spoke being echoed back to them across miles of scenery, two people fundamentally disagreeing over what verbal information they had just heard spoken, these situations are creepy and hilarious; mundane or even startling, and undeniably interesting.

In regards to human history, visual and tactile mediums had long been the only formality by which to reliably record and document the present tense moment (art, writing, etc.). In the sonic sense, information has deeply relied upon notation and our use of verbal language through storytelling and music passed down orally through generations. Sonic mediums that allow people to reliably record information did not exist for much of our history, and only first became available with the invention of the phonograph by Thomas Edison in 1877 (Thompson, 2016). The progression of technology and the invention of the internet have had a massive cultural impact on our relationship with sound. Cell phones have made it so that music can easily be heard and taken with us everywhere; recording software has become so compact that it can fit into one's pocket. The recent uptick in studies exploring auditory bistability has a lot to do with the progression of social media tech giants such as Twitter and Instagram, having made sharing any perceptually baffling information into a potentially viral event. Many different auditory bistable stimuli have made their way to internet infamy by being shared across these platforms (Gajanan, 2018).

In May of 2018, a short audio clip took the internet by storm, causing a multitude of confusion with its perceptual bistability. The stimulus is often referred to as the name of both mutually exclusive stable states it is commonly perceived as: "Laurel & Yanny". The sound clip was originally recorded by an opera singer and actor named Jay Aubrey Jones, who was contracted by the founder of "vocabulary.com" to record English words for the website (Gajanan, 2018). It was then posted to Reddit by a high school student in Georgia named Roland Szabo who pointed out that it confusingly sounded as though someone were saying "Yanny"; and shortly after it became internet famous after being posted to Twitter by the account "CloeCouture" (Gajanan, 2018).

This bistable stimulus presents a unique opportunity to explore auditory bistability in a short auditory stimulus (just 670 milliseconds long). The study of auditory bistability had remained to be largely uncharted territory as compared to the amount of literature that examined visual perception and visual bistability. In recent years, psychologists and cognitive neuroscientists have begun to explore bistability much more, however these bistable stimuli differ from "Laurel & Yanny". For one, the majority of auditory stimuli explored previously are auditory streaming tones that last continuously for long periods (Davidson & Pitts, 2014; Higgins et al., 2020). "Laurel & Yanny" differs from previous examples of auditory bistability in that it is an actual human voice uttering an abrupt; two-syllable spoken word. This stimulus thus presents an opportunity to explore brain activity that occurs when experiencing perceptual reversals in a one-word verbal language stimulus. To listen to this stimulus for yourself, visit the following YouTube link (Guardian News, 2018):

[https://www.youtube.com/watch?v=7X_WvGAhMlQ]

What do you perceive? Can you switch between hearing "Laurel" and "Yanny" as you listen to repeated presentations of the sound clip?

1.2.3 Theories Behind Bistability

Information processing is a theoretical approach to behavior and perceptual awareness which attempts to define sensory processing based on a series of cognitive processes (Ward, 2020). There are many proposed explanations that attempt to define the processes behind perceptual bistability, which fall into two main categorical strategies of information processing (Intaite et al., 2010). The first category is referred to as "bottom-up" processing, which can be defined as the gathering of sensory information, the assemblance of this information, and the integration of this information into higher order cognitive processes. This can more simply be thought of as the "passage of information from simpler to more complex" (Ward, 2020, p. 7) starting with the peripheral sensory receptors and transmitting up through the levels of subcortical and cortical processing, or as the sensory question "What am I currently witnessing?" These theories behind perceptual bistability offer the idea that low-level sensory mechanisms are to blame. They propose that two competing perceptual mechanisms exist when experiencing a bistable stimulus: one of which dominates over the other as we continuously perceive one stable state. However, once this dominant mechanism gets "worn out", the secondary mechanism dominates our perceptibility and we perceive the secondary stable state of this bistable stimulus as the result (Cohen, 1959). It has been theorized that as long as both stable states are equivalently likely to take place; one will experience a perceptual reversal on a bistable stimulus (Kornmeier & Bach, 2005).

This concept of a population of neurons becoming "worn out" for an amount of time specifically refers to synaptic fatigue. Using a bistable stimulus in our example, this fatigue has been theorized to occur due to high frequency stimulation of the currently dominant neuronal population. Supposing this neuronal population influences the perception of "Laurel", its overstimulation would temporarily lead to the depletion of the number of neurotransmitter vesicles capable of binding to and fusing with the plasma membrane of this population's presynaptic neurons (Luo, 2015). This would thereby inhibit transmitters from firing into the synaptic cleft (the junction between two neurons). Another neuronal population would then take over; the secondary population supposedly influencing this new perception of the stimulus: “Yanny”.

The other category that many common explanations for perceptual bistability fall into is referred to as “top-down” processing. Top-down processing can be defined as the process whereby our expectations, prior knowledge, and higher order cognition influences our predictions and interpretations of the sensory information we are experiencing. This category is in many ways inverse to that of bottom-up processing, and can be more simply thought of as “the influence of later stages on the processing of earlier ones” (Ward, 2020, p. 7); for example: the prefrontal cortex or the hippocampus influencing the lower-level sensory cortex (Xu et al., 2020). Top-down processing can also be thought of as the conceptual question: “Have I witnessed this before?” These theories behind perceptual bistability propose that perceptual reversals are due to the feedback regulation of lower level processes by central (higher-level) processing mechanisms (Girgus, Rock, & Egatz, 1977; Intaite et al., 2010).

These two categorical theories behind perceptual bistability have not been proven to be mutually exclusive, and some influential theories of bistable perception explicitly propose a “hybrid” approach to understanding the interacting dynamics between bottom-up and top-down mechanisms (Long & Toppino, 2004). Indeed, it remains plausible that bottom-up and top-down processes may come to influence one another in some form; for example it is possible that negative feedback could influence synaptic fatigue (Intaite et al., 2010; Luo, 2015). Many of these theories have been developed through research involving electrophysiological recordings of the brain, known as electroencephalography (EEG); with other brain imaging techniques also contributing to the discussion. EEG can help us find neural markers that occur when participants experience a given stimuli or a mental event, such as a change in perception.

1.3 EEG & ERPs: Electroencephalography & Event-Related Potentials

One method that allows us to study the relationship between brain activity and perception is known as EEG. When the EEG is recorded from the scalp, it is one of many methods of functional imaging that are non-invasive, meaning that it is harmless to the body and does not require physical infiltration to record brain activity. Instead, participants simply wear a cap that is tightly fitted to their scalp. Changes in voltage signals across different populations of neurons are measured temporally (across time) by metallic electrodes. These electrodes are allocated across this cloth cap, and record the summation of postsynaptic potentials emitted by neurons in the brains of each participant.

Neurons fire at lightning-quick millisecond speed, delivering an extremely brief active current (just 1 millisecond long!) traversing the axonal length of the neuron and a secondary passive current which runs across the dendrites and the nucleus of the following neuron (Ward, 2020). It is these passive postsynaptic currents which are measured by EEG, due to their slightly longer temporal duration; lasting tens to hundreds of milliseconds long rather than for one brief millisecond (Nunez, 1981). Luckily for EEG, neurons in the cerebral cortex (closest to the scalp) are mostly arranged in a parallel orientation. This orientation allows them to summate together and thus to not be canceled out by one another's electrical currents, helping us to reliably record neuronal transmission within the cerebral cortex as a result (as well as other regions of the brain that are laid out with such organization; Ward, 2020). Firing synchronicity between two neurons is theorized to be indicative of neuronal communication (Fries, 2005). For EEG to reliably detect nerve cell firings at the scalp level; groups of neurons must generate a large enough postsynaptic electrical current which occurs as a direct result of neuronal populations firing in synchronicity with one another (Ward, 2020).

Like any functional imaging method used to record brain activity, EEG has its strengths and its weaknesses. Where EEG falls short is in its spatial resolution (locational accuracy). As one might think, attempting to record brain activity at scalp locations does not give one the tools necessary to infer where neuronal populations or pathways of interest may specifically lie within the brain itself. In addition to this, not every region of the brain has neural populations oriented in parallel, making it impossible to record brain regions like the thalamus as a direct result of this region's non-parallel neuronal orientation (Ward, 2020). EEG is instead well known and revered for its striking temporal resolution (timing accuracy); being able to precisely record the sequence of neural activity (at a millisecond resolution) that occur in response to sensory stimulation and during certain tasks in an experimental paradigm.

To accurately assess data and measurements of brain activity that occur during a specified task, experimenters must reliably coordinate EEG recording in real time to an externally observable event that can be precisely time-stamped. Doing this is crucial, as specific brain activity patterns linked with a particular sensory, cognitive, or motor process are hidden within the larger "background EEG". Extracting these event-related potentials (ERPs) from the background EEG is accomplished via the process of "time-locking" the recording to the events of interest (e.g., stimulus onset), and averaging the data across many (hundreds) of trials. In the case of bistable perception, if we wish to track the moment one experiences a perceptual reversal from one stable state of a stimulus to another; we can have participants press a button each time we present the stimulus for them to report how they perceive it. "Trigger codes" indicating when the stimulus was presented and which button the subject pressed are then time stamped within the EEG data, letting us know exactly when these instances occurred. We can then look at successive trials and sort them according to the button presses into "reversal" versus "stable" trials, as well as "Laurel" versus "Yanny" trials.

1.4 ERP Neural Correlates of Bistable Perception

1.4.1 The Reversal Negativity

The Reversal Negativity (RN) is an ERP component that was discovered in 2004 by Kornmeier & Bach. It has since been documented across multiple studies using visual bistable figures (Abdallah & Brooks, 2020 ; Britz et al., 2009; Britz & Pitts, 2011 ; Intaite et al., 2010 ; Kornmeier & Bach, 2004, 2005; Kornmeier et al., 2007; Pitts, 2007, 2009). The RN manifests as a relatively negative-going amplitude shift for reversal versus stable trials from ~150-350 milliseconds post stimulus onset, bilaterally over the occipital-parietal scalp. In two EEG studies, an auditory reversal negativity (aRN) has been observed (Davidson & Pitts, 2014; Higgins et al., 2020). Both studies used bistable auditory streaming sounds constructed of multiple tones. The streaming stimuli presented in Davidson & Pitt's (2014) study were Shepard's tones which elicited the perceptual sensation of either ascending or descending pitch motion, while the stimuli presented in Higgins et al (2020) were triplets of ABA tones followed by a brief tone absence; dually perceivable as either one "galloping" stream or as two separate streams. In both experiments, stimuli were presented intermittently while participants responded to each presentation in accordance with their perception.

Davidson & Pitts (2014) located an aRN at a slightly earlier onset time and with a more frontal central scalp distribution. Higgins et al (2020) located significantly more negative ERPs for perceptual reversals, with a later onset timing and a similar frontal central scalp distribution. These reversal negativities are theorized to be neural markers indexing the transition from one mutually exclusive state of a bistable percept to the other state.

1.4.2 The Late Positive Complex

The late positive complex (LPC) is an ERP speculated to be associated with post-perceptual processing of a perceived reversal. It is speculated to be related to the retrieval of one's subjective perception in order to report one's perceptual experience, typically occurring roughly ~300 - 400 milliseconds post-stimulus at central-parietal locations of the brain, similar to the well-known P3b wave (Dien & Donchin, 2004). The LPC has been found in studies that require subjects to report perceptual switches (or non-switches) as well as studies that task subjects with reporting individual percepts on each trial, e.g., face or vase (Dien & Donchin, 2004; Davidson & Pitts, 2014). Thus, the LPC does not seem to be related to only reporting perceptual switches. Rather, the LPC seems related to deciding whether to respond, or how to respond; which may change from trial to trial based on the participant's perception (Yang et al., 2019). As noted in a 2019 study by Yang et al., the LPC is task-dependent: "It was present only when the memory judgement at hand required consideration of the relevant dimension" (Yang et al., 2019, discussion par. 1). Davidson & Pitts (2014) found that in regard to auditory bistability, the aLPC (auditory LPC) occurred at a correspondingly similar temporal latency to that of the visual LPC, with a relatively similar scalp topography (more fronto-centrally); appearing 350 milliseconds post-stimulus.

1.5 The Present Experiment: Summary, Question & Hypothesis

The present study aims to investigate neural markers of bistability that occur in response to the verbal audio clip known as "Laurel & Yanny"; perceived by most listeners as either one of the former names for a period of time. There are a few experimental questions and goals we wish to investigate using this auditory bistable stimulus. First of all, across successive presentations of the stimulus; is it possible to detect a neural difference (measured via EEG/ERPs) when subjects indicate a shift in their perception from Laurel-to-Yanny or Yanny-to-Laurel? How might this compare to perceiving the same name across successive trials (i.e.; Laurel-to-Laurel or Yanny-to-Yanny)? This contrast can serve to tell us about the neural events involved in changes in conscious perception. Preliminary evidence suggests this perceptual reversal should produce an "auditory reversal negativity" (aRN), but only two studies have identified such an auditory reversal negativity to date (Davidson & Pitts, 2014; Higgins et al., 2020).

In the event that an aRN and LPC are found, a secondary goal of the present study will be to question whether there is any correlation between participant behavior (reversal rate) and the amplitude (μV) of each electrophysiological effect. This will be done in order to identify if the effect in question may be an oddball effect like the mismatch negativity (MMN); which is "caused by the brain's automatic reaction to unexpected changes in a repetitive stimulation" (Sandberg et al., 2005, abstract par. 1). Thus, any statistically significant correlation between participant’s reversal rates and an electrophysiological component’s amplitude will reflect an oddball effect, and if the correlated ERP is negative-going and occurs between ~100 – 200 milliseconds; it would specifically suggest an MMN-like component. Likewise, no correlation between reversal rate and an effect’s amplitude would signify that the electrophysiological component in question would be interpretable as a neural correlate of bistable perception.

A final question we wish to explore is to determine: is it possible to detect a neural difference (measured via EEG/ERPs) when subjects report perceiving either Laurel (Percept A) or Yanny (Percept B)? If this be the case; when will this neural difference occur temporally (in relation to the onset of the stimuli) and where on the scalp will the neural difference occur spatially (in terms of topographical distribution of the ERP effect)?

The contrast between perceiving “Laurel” and “Yanny'' may give us a clue about when in the stages of sensory processing perception begins to part ways from the encoding of the physical world. This data may serve to help us pick apart details of how or when such stages of sensory processing occur, lending us details regarding a shift to a mental understanding of the physical world.

Chapter 2: Methods

2.1 Participants

31 participants with no documented history of neurological conditions, brain damage or deafness were recruited for this experiment (12 male, 19 female, 18-22 years of age, mean age = 20). Participants were paid \$30 for their participation in this study; 10\$ for participation in the pretest and 20\$ for participation in the experiment. We received written informed consent from each participant at the beginning of each experimental session. All procedures in this experiment were approved by the Reed College Institutional Review Board and adhere to federal ethics regulations.

2.2 Stimulus: Laurel & Yanny

The stimulus was downloaded from a YouTube repost of the original source audio file (originally posted on Twitter). One iteration of the stimulus lasts approximately 670 milliseconds. The audio file was uploaded to Logic Pro X for auditory manipulation. A total of eleven stimulus variants were created, each of which were of different semitone tunings. Stimuli were manipulated via the "Pitch Shifter" plugin, and each variant was programmed to be one semitone apart: creating a spectrum of 11 stimuli. A waveform image is depicted in Figure 2.1.

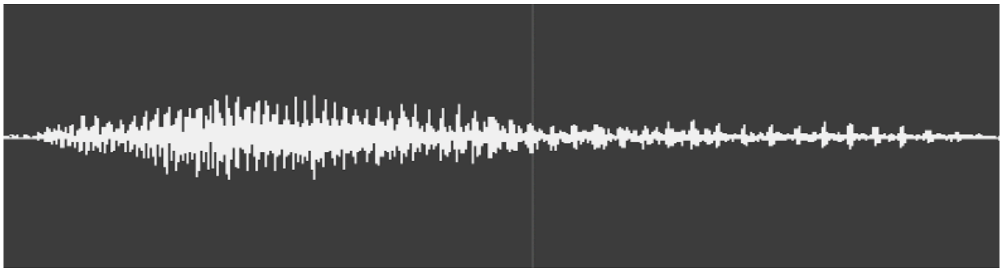


Figure 2.1 "Laurel & Yanny" Waveform. ~670ms long in duration, the internet famous stimulus can be heard as either the name "Laurel" or the pseudonym "Yanny".

In a pilot experiment, all eleven variants of the stimulus were played in consecutive order 25 times. Participants (N = 11) reported when they experienced perceptual reversals upon listening to each variant of the auditory stimulus. Based on these results, we chose the stimulus with the highest number of reversals to be the new central stimulus in the semitone adjustment spectrum. Pilot experiment results revealed that participants reported perceiving “Yanny” as the semitone tuning of the stimulus became increasingly negative and “Laurel” as the semitone tuning of the stimulus became increasingly positive. A new spectrum of eleven stimulus variants was created for use in the pretest based on these pilot results: the most “Yanny” biased stimulus variant at a semitone frequency of -7, the central stimulus at a semitone frequency of -2, and the most “Laurel” biased stimulus variant at a semitone frequency of +3. A graphic depicting this spectrum of eleven stimulus variants can be found in Figure 2.2.

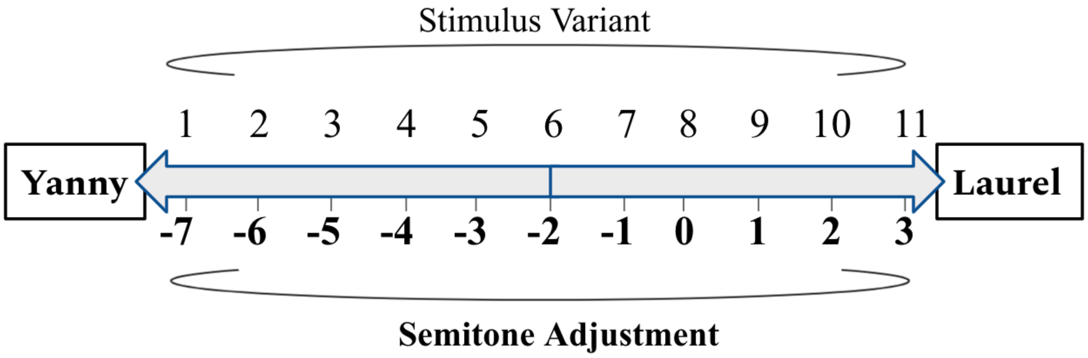


Figure 2.2 “Laurel & Yanny” Pretest Stimuli Spectrum. Eleven stimulus variants were designed for the pretest of the experiment. All were tuned one semitone interval apart from one other. The top row above the horizontal arrow depicts the numeric label assigned to each stimulus variant, while the bottom row depicts the semitone tuning of each variant (relative to the original stimulus downloaded from YouTube). As semitone pitch became higher, pilot results revealed participants perceived “Laurel” more often. As semitone pitch became lower, pilot results revealed participants perceived “Yanny” more often. The central stimulus was pitched to -2 semitones, with the far end of Laurel pitched down -7 semitones and the far end of Yanny pitched up +3 semitones. The original unaltered stimulus (variant 8) was positioned right of center, interpretable as “Laurel” by the majority of pilot participants.

All stimuli were bounced from Logic Pro X as WAV files, at 16-bit resolution and at a sample rate of 44100. Stimuli were uploaded and presented to participants using Presentation (Neurobehavioral Systems, Berkeley CA) at approximately 80 decibels of volume via two (stereo) THX certified Z906 Logitech Surround Sound Speakers (frequency response: 35 Hz–20 kHz). Speakers were placed roughly 2-3 feet away from participants, 100 degrees apart from each other and positioned inward facing participants to their left and right.

2.3 Procedure

The study consisted of two parts: the pretest and the actual experiment. In both cases, participants were instructed to press one of two keys with their right hand after listening to each stimulus. Participants were asked to press the “left arrow key” with their pointer finger if they heard “Laurel”, and to press the “down arrow key” with their middle finger if they heard “Yanny” (response keys were counterbalanced across participants). If they were unsure of what they heard or if they heard a superimposed mixture of Laurel-and-Yanny, they were instructed not to press anything, and such trials were discarded from analysis as it was not an objective of the study to compare trials in which participants were unsure of their perceptual experience.

For both the pretest and actual experiment, participants were placed inside a sound attenuated room with the previously described speakers set up for stimulus presentation. Upon first sitting in the room for the pretest, the central stimulus (variant 6) was presented to the participant previous to any experimental conditions to accustom them to both the volume (~80 dB) and the bistable nature of the stimulus. Participants were asked whether they had ever heard the stimulus before after it was played for them.

The room was moderately lit by a small overhead ceiling lamp. Participants were asked to keep their focus on a fixation dot in the center of a monitor placed before them so as not to disrupt EEG data with any unwanted eye movement or muscle-tension related artifacts. During all trials of the pretest and of the experiment, the ~670ms stimulus was played for participants followed by an inter-stimulus silent interval that jittered within a random distribution from 700 to 900 milliseconds: 800 milliseconds on average. A graphic depicting this is provided in Figure 2.3.

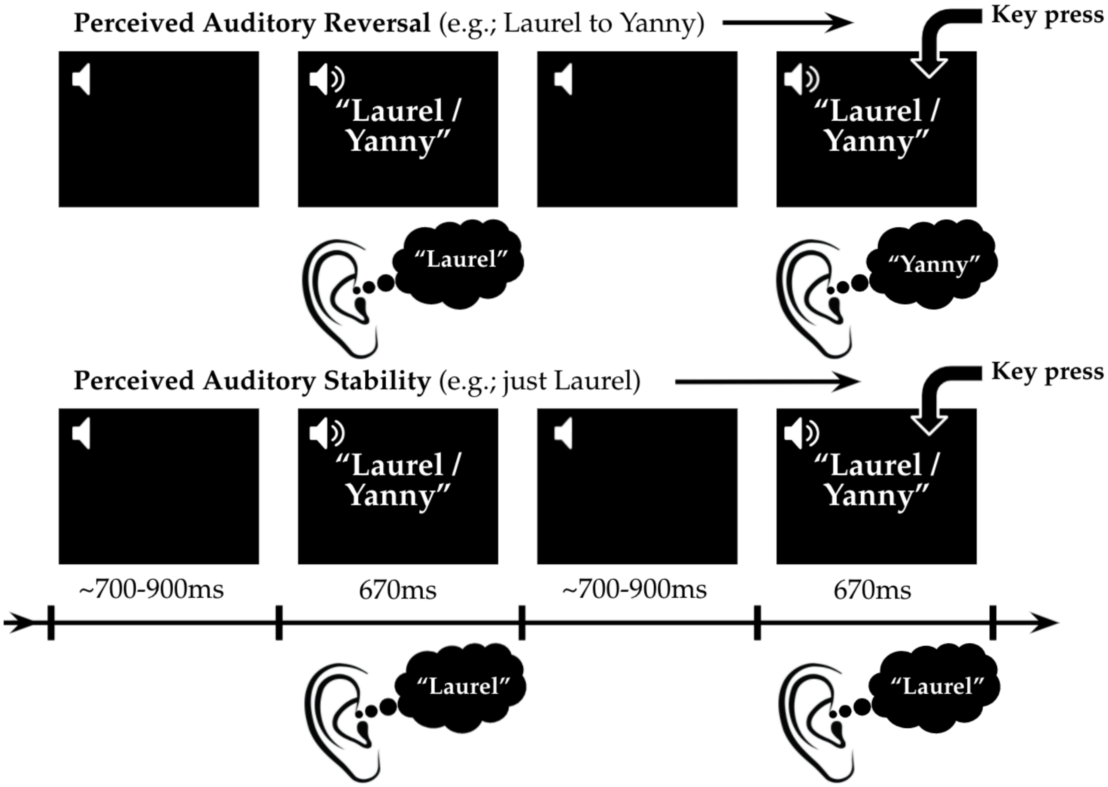


Figure 2.3 Experimental Procedure.

Participants sat and listened to a randomly jittering ~700 - 900ms silent interval, and then the bistable auditory stimulus "Laurel & Yanny" for ~670ms. Participants then pressed the corresponding "Laurel" key or "Yanny" key with their right hand depending upon their perception of the word. A perceived auditory reversal trial from "Laurel" to "Yanny" is shown on the top row. A perceived auditory stability trial from "Laurel" to the same percept "Laurel" is shown on the bottom row. Both types of reversals (Laurel-to-Yanny and Yanny-to-Laurel) were collapsed for analysis of "perceptual reversals", and both types of stable trials (Laurel-to-Laurel and Yanny-to-Yanny) were collapsed for analysis of "perceptual stability".

The original objective of the pretest condition was to assess the stimulus that participants reported experiencing the most reversals upon hearing, with the intention of using this stimulus for all blocks of the actual experiment. During the pretest, participants listened to each of the eleven semitone variants of the stimulus, consecutively 25 times in a row. They reported their perception during each presentation of the stimulus. The same inter-stimulus interval used during experimental blocks was implemented. Stimulus variants were designed to play in a recurring order during the pretest: two biased forms of the stimulus were played, one variant biased towards Laurel in semitone frequency and the other biased towards Yanny in semitone frequency; followed by a more ambiguous semitone variant. The precise order that stimulus variants were played for participants during the pretest was: 2_{-6}, 10_{2}, 5_{-3}, 7_{-1}, 11_{3}, 3_{-5}, 6_{-2}, 8_{0}, 1_{-7}, 9_{1}, and 4_{-4} (subscripts denote their semitone tuning difference from the original stimulus).

After participants had listened to all eleven stimuli, the data from their responses was quickly uploaded and run through a MATLAB plotting script to generate an informative graphic of their responses. The stimulus that the participant experienced the most perceptual reversals on was then used for the actual experiment. The entire pretest on average took about 5 to 7 minutes to complete.

For the actual experiment, participants were fitted to EEG caps. The experiment initially consisted of 10 blocks of 80 trials, which amounted to 800 listened presentations of the stimulus. Between blocks of the experiment (as well as the pretest), participants were offered the ability to rest their eyes and ears, and to take a quick break if they wanted to; seated in the experiment room. The experiment lasted approximately 20 minutes on average. In total, both the pretest and the experiment lasted between 30 to 40 minutes on average. The EEG capping and experimental prep took anywhere from 30 minutes to 1 hour to reliably complete, resulting in a total experimental session time of ~1.5 hours per subject.

Upon realizing that certain participants were not reversing their perception of the stimulus at a fast enough rate to result in enough trial segments for inclusion in the reversal ERPs (N = 4); additional steps were taken to improve the pre-screening and training of subjects prior to the main experiment.

Participants were emailed instructions detailing a short 5-min practice with the stimulus, and were told to inform the lab ahead of time if they could not hear both names or if they had difficulty reversing in perception at a reasonably quick rate. The pretest also became a dry-run (non-EEG) screening test in which subjects needed to switch on a stimulus in order to be eligible for participation in the full EEG experiment. Once a participant entered the lab, they were informed of the task and subsequently sat in the experiment booth to do the pretest and ensure they were eligible for actual experiment blocks. If a participant could not experience perceptual reversals, their pay remained at 10$ for their time running the pretest. The original functional goal of the pretest (to determine which stimulus variant one could perceptually reverse on the most) remained in effect.

To further prevent the outcome of having too few reversal trials for EEG: two more blocks were added to the main experiment, bringing the standard block total up to 12 and the trial total up to 960. Additional blocks were added for some participants (N = 3) who did not experience enough reversals during the initial 12 blocks.

2.4 EEG Recording

To record brain activity during the experiment, participants were fitted to a 64-channel electrode EEG “EasyCap”. In order to reliably detect and remove motor artifacts resulting from eye movement and eye blinks, three electro-oculogram (EOG) electrodes were used. One was placed vertically below the left eye (VEOG) for eye blink and up-down eye motion detection while the other two were placed horizontally to the left and right eyes (HEOG) for left-right eye motion detection. For all electrodes impedance levels were kept below 5kΩ; achieved via the implementation of wooden Q-tip scalp abrasion and placement of saline-gel within each electrode. EEG was recorded via the recording application Brain-Vision Recorder (Brain Products, Gilching Germany). An image depicting the model of the EEG cap and the locations of all 64 electrodes is provided in Figure 2.4.

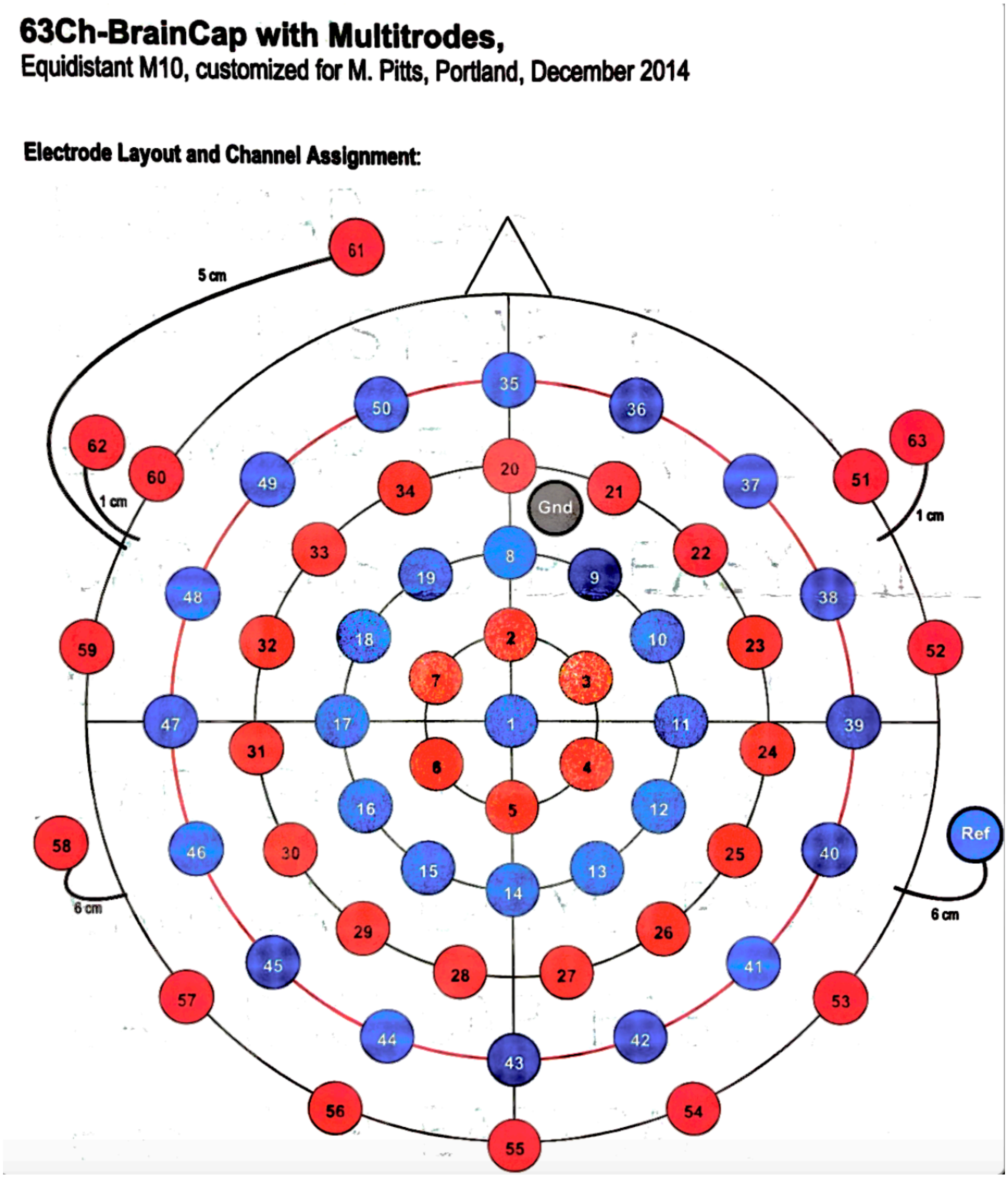


Figure 2.4 64-channel electrode EEG “EasyCap” (Herrsching, Germany) used to record brain activity during experimental blocks. The right mastoid electrode (reference electrode “Ref” seen in blue) and left mastoid electrode 58 (in red) are bilaterally visible just below the horizontal diameter of the cap near the ears. The left sub-ocular VEOG electrode 61 is visible anteriorly just to the left of the nose. The HEOG electrodes 62 and 63 are visible bilaterally at the periphery of each eye.

2.5 Data Analysis

Brain-Vision Analyzer (Brain Products, Gilching Germany) was used to pre-process and analyze the EEG data. Preprocessing pipeline steps consisted of segmentation for the selected condition (either Laurel, Yanny, Reversal or Stable), epoching (-200ms to +800ms relative to stimulus onset), filtering (25Hz low-pass, 12dB roll-off), re-referencing (average mastoid referencing for all scalp channels, bipolar referencing for VEOG and HEOG), artifact detection/rejection (blinks, eye movement and C.R.A.P. "Commonly Rejected Artifactual Potentials"), averaging, and baseline correction (-200ms to 0ms); finally resulting in the creation of mastoid reference ERPs and average reference ERPs for the selected condition (e.g.; Yanny ERPs). Independent component analysis (ICA) was performed in order to detect and correct (rather than reject) eye-blinks and eye-motion artifacts from the EEG data of a few participants (N = 2) who had a low number of trials remaining in at least one condition post-artifact-rejection (<70 trials).

Pretest data was analyzed by adding up all uses of the eleven stimuli variants. This was done in order to determine which semitone tunings of "Laurel & Yanny" were the most bistable for participants. Of the 31 total subjects that completed the pre-test, 24 qualified for the main experiment (77%). The behavioral data analysis included all participants who participated in the main EEG study (N = 24), including those that were later excluded (N = 3) due to excessive EEG artifacts or partially excluded (N = 2) due to insufficient trial numbers in one of the main conditions (<70 trials).

In order to assess how easily participants could perceive both names, a myriad of behavioral data analyses were conducted. For all participants, the percentage that each perceived Yanny and Laurel was calculated; in addition to an entire dataset mean value for Yanny percent perceived and Laurel percent perceived across all subjects. Similarly, the percentage of stable trials perceived and reversal trials that occurred was calculated for each participant, as well as for the entire data set to generate a mean value for percent Stable trials perceived and percent Reversal trials perceived across all subjects.

Reversal rates were also calculated from participant's responses. One of the most crucial behavioral data calculations generated, reversal rates indicated how many stable trials on average there were before each switch in perception. This was calculated for each participant individually and was also calculated as a mean across all participants whose data were used for all conditions of the study. This calculation can best be explained as the average number of stimuli before a reversal occurs, which was quantified by dividing the number of stimuli between reversals during a block by the number of reversals in a block; and then averaging these values out across the total number of blocks a participant sat through in their experimental session.

To statistically analyze electrophysiological effects, difference waves of reversal trials minus stable trials were generated across all participants and analyzed via a false discovery rate (FDR) mass univariate analysis (MUA) in MATLAB, using the MUA ERP toolbox (Groppe et al., 2011). The same analysis (FDR MUA) was performed with Laurel minus Yanny difference waves in an effort to detect any neural differences between the perception of the two names.

We also investigated how electrophysiological effects compared to behavioral data in an effort to answer the question: Do reversal rates predict the amplitude of any electrophysiological effects? Put differently, does the amplitude of any electrophysiological effect found correlate with an increase in the number of stable trials before a reversal? For this analysis, mean amplitudes for each electrophysiological effect found in the MUA were calculated for each electrode cluster and time-window where the amplitude difference was largest. These mean amplitudes were calculated for all participants individually for all electrophysiological effects found. Participant's mean amplitudes for each effect were then compared to their reversal rate to determine if a correlation existed.

If so; electrophysiological effects that do correlate with amplitudes may be considered an oddball effect; much like the mismatch negativity (MMN). In this case, the more stable trials before a reversal rate; the larger the amplitude of the effect. If the reversal rates do not predict the amplitude of the effects, then any electrophysiological effects found are more validly interpretable as neural correlates indexing changes in conscious perception.

Chapter 3: Results

3.1 Behavioral Results

3.1.1 Pretest Results

Of the total participants (N = 24) who passed the pretest and participated in the full experiment (including those excluded from the EEG analysis), variant 8 was the most bistable variant in the study (33.33%); which was tuned down -3 semitones from the original stimulus pitch and selected by 8 participants during the pretest. Stimulus variant 7 was the second most bistable variant in the study, which was tuned down -1 semitone from the original stimulus pitch and selected by 5 participants (20.83%) during the pretest. Stimulus variants 4 and 6 tied for third place, each being perceived by N = 3 participants (12.5%) to be the most bistable variant (the former tuned down -4 semitones and the latter -2 semitones).

Variant 8 was the original unedited stimulus and was only used once in the entire study. Variants 3 and 11 were also perceived to be bistable by a participant (N = 1 each). Variants 1, 9, and 10 were not selected by any participant during the pretest and thus were not used in the study. A plot depicting stimulus variants selected during the pretest and used in the experiment can be viewed in Figure 3.1.

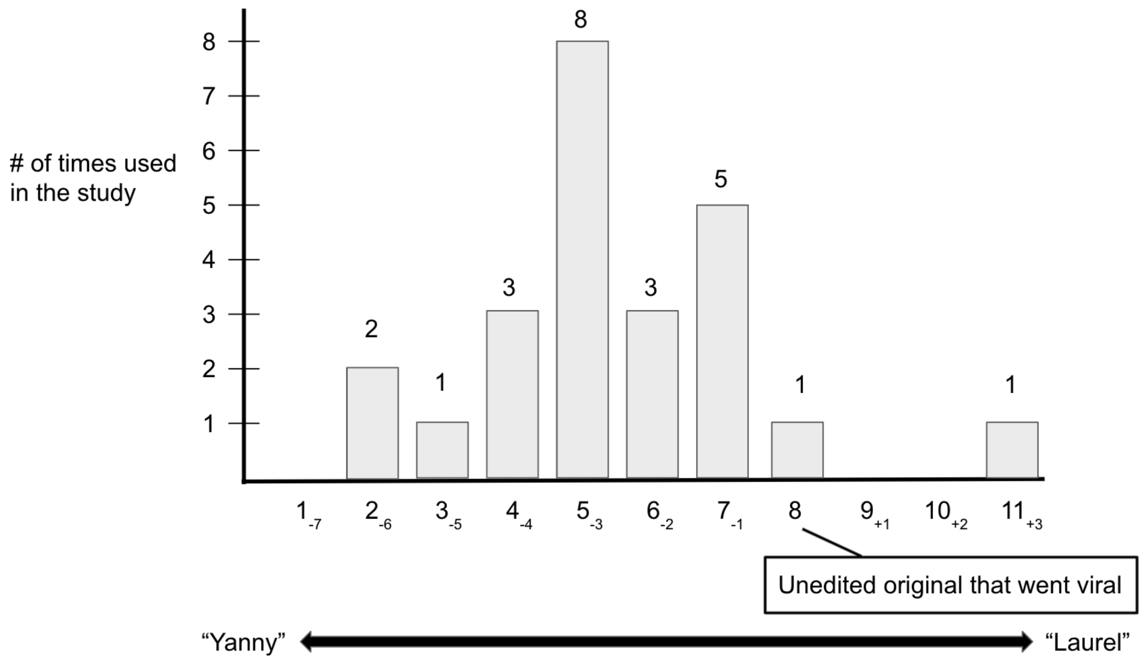


Figure 3.1 Pretest Stimuli Distribution.

Bar plot depicting the number of times each stimulus variant was used in the study by a participant. Plotted on the horizontal axis are all eleven "Laurel & Yanny" stimulus variants, each subscript containing their respective semitone tuning. The number of times the stimulus was selected by a participant during the pretest (and thereby used in actual experiment blocks) is reflected on the vertical axis. The central stimulus was pitched to -2 semitones, with the far end of Laurel (variant 11) pitched down -7 semitones and the far end of Yanny (variant 1) pitched up +3 semitones. The original unaltered stimulus (variant 8) was positioned right of center due to it being perceived as "Laurel" by the majority of pilot experiment participants.

3.1.2 Participant Exclusion

Of 31 initially recruited participants, N = 7 did not pass the pretest and thus were not eligible to participate in the study. N = 24 thus participated in the full EEG experiment.

In order for a participant to be included in the main EEG analyses, they must have obtained a minimum of 70 reversal trials post-artifact rejection. Preliminary data from the first group of EEG participants suggested that some subjects were not yielding enough reversal trials (N = 4), some due to artifact-rejection of reversal trials (of the aforementioned; N = 2). Although some subjects did experience perceptual reversals they did so at an incredibly slow rate; perceiving “Yanny” at times for ~40 trials in a given block and then switching in perception to “Laurel” for the next ~40 trials (N = 2). This low rate of perceptual reversals caused there to be a great magnitude of stable trials and a comparably low number of reversal trials in their EEG data. Due to this uneven distribution and its effects on signal-to-noise ratios in the ERP data, it would have been invalid to compare their high number of stable ERPs to their low number of reversal ERPs; thus, they were ineligible for the reversal-stable analysis. Two previously mentioned participants exhibited plentiful numbers of stable trials consisting of both percepts, thus being eligible for inclusion in the Laurel-Yanny analysis.

A few participants became eligible for the reversal-stable analysis due to independent component analysis (ICA) which helped to salvage reversal trials (N = 2). Blocks were also added if participants did not reverse in perception enough during the initial 12 blocks of the experiment (N = 3) in an attempt to raise the amount of reversal trials usable for the planned analysis. One participant was excluded from the study due to failure to correctly execute the task.

Participant sample sizes for the electrophysiological analyses:
In total, 19 participants yielded EEG data applicable for all planned analyses of the experiment: the reversal-stable analysis and the Laurel-Yanny analysis. 21 participants yielded EEG data applicable for the Laurel-Yanny planned analysis.

3.1.3 Stimulus Perception

Averaged across all participant's experiment trials (pre-artifact rejection), both "Laurel" and "Yanny" were equally perceived; "Laurel" being perceived 49.82% of the time (M = 49.82% ; SD = 10.30 ; E = 2.36), "Yanny" being perceived 50.17% of the time (M = 50.17% ; SD = 10.30 ; E = 2.36). Max-min values for the percentage that "Laurel" was perceived were 65.39% & 27.3% (Range = 47.53). Max-min values for the percentage that "Yanny" was perceived were 72.69% & 34.78% (Range = 38.07).

Averaged across all participant's experiment trials (pre-artifact rejection), "Stable" trials were perceived 3.17 times more than "Reversal" trials were: "Stable" trials were perceived 73% of the time (M = 73% ; SD = 16.52 ; E = 3.79) while "Reversal" trials were perceived 27% of the time (M = 27% ; SD = 16.52 ; E = 3.79). Max-min values for the percentage of "Stable" trials perceived were 91.4% & 19.96% (Max-Min = 91.4 - 19.96, Range = 71.44). Max-min values for the percentage of "Reversal" trials perceived were 80.04 - 8.6 (Max-Min: 80.04% - 8.6% , Range = 71.44).

A table depicting the average percentage in which stable, reversal, "Laurel", and "Yanny" trials were perceived is provided in Figure A1 of the Appendix.

3.1.4 Reversal Rate

Reversal rates were computed as the average number of stable trials that took place before a reversal occurred. Thus, a larger reversal rate means that it took longer to perceive a perceptual reversal, whereas a smaller reversal rate reflects that it took shorter to perceive a perceptual reversal. The mean reversal rate for total participants was M = 6.8 stimuli (SD = 3.55 ; E = 0.81). Max-min values were 16.9 & 3.18. Reversal rates were quantified for each participant by dividing the number of stable trials between reversals by the number of reversals within each block. These values were then averaged out across the total blocks for each participant. Stable trials at the end of a block were excluded from calculations as it was not possible to determine when a participant would have switched next. Figure 3.2 shows a frequency plot depicting the distribution of reversal rates across participants. Figure A2 of the Appendix reports reversal rates for each individual participant and the average reversal rate across total participants.

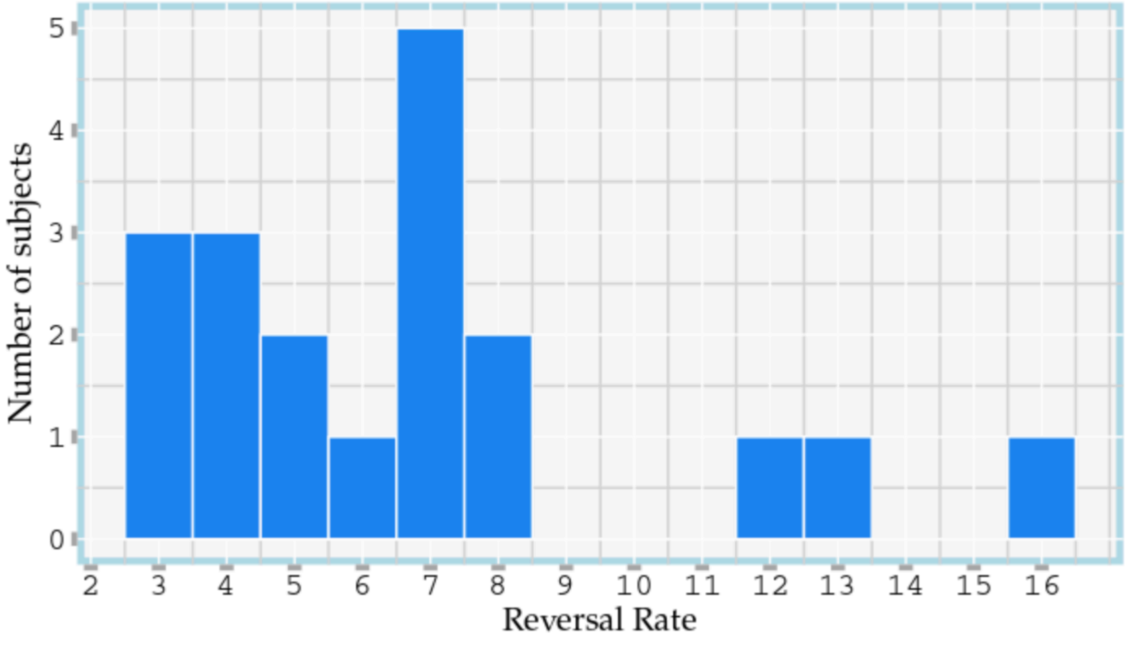


Figure 3.2 Reversal Rate Frequency Plot. Distribution of reversal rates across participants analyzed for all conditions (N = 19). Reversal rates were quantified as the average percentage of stable trials that took place before a reversal occurred. The mean reversal rate for total participants was M = 6.8 (approximately 7 stable trials before each perceptual reversal).

3.2 Electrophysiological Results

3.2.1 Reversal vs. Stable: Mass Univariate Analysis

Difference waves of participant's reversal minus stable trials were analyzed within a false discovery rate mass univariate analysis, revealing three effects: two negativities and one late positive complex. The first negativity was evident from ~120 to 180 milliseconds post-stimulus, encompassing central and parietal electrodes. The second negativity was evident from ~200 to 350 milliseconds post-stimulus, concentrated bilaterally on temporal locations with a left-hemisphere bias. The third effect manifested as a sizable late positive complex, occurring ~400 to 1000 milliseconds post-stimulus and detected centro-parietally. The mass univariate results are shown in Figure 3.3.

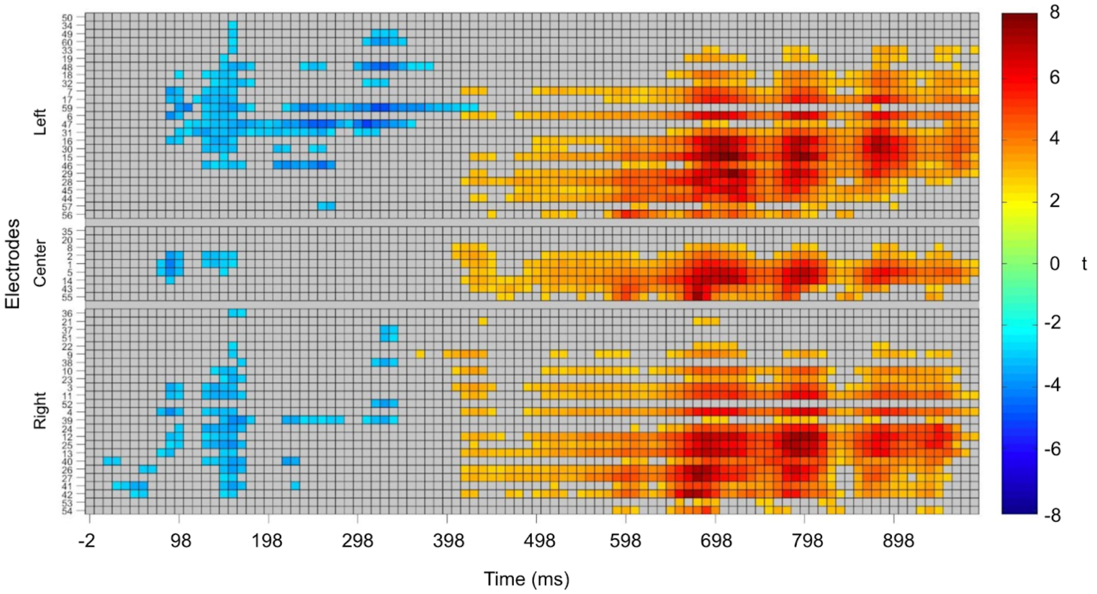


Figure 3.3 Reversal minus Stable false discovery rate mass univariate analysis. Electrode locations are plotted over time (ms) post-stimulus onset. On the vertical axis are electrodes. From the top down are three sets of rows, the first outlining left hemisphere electrodes, the second midline electrodes, and the third right hemisphere electrodes. Within each set of rows, electrodes are arranged from anterior to posterior. The horizontal color-bar to the right of the chart reflects corresponding t-scores for all amplitude differences (5% false positive rate, corrected for multiple comparisons). Light and dark blue indicate negative amplitude differences that were significant while orange and red indicate positive amplitude differences that were significant.

3.2.2 First Negativity

An early negativity was evident in the mass univariate analysis from ~120 - 180 milliseconds post-stimulus, detected upon central and right parietal locations. Central electrode 17 (electrode C3 on the international standard 10-20 system) detected the largest negative amplitude difference (μV) at this time window, marked in yellow on the left-facing voltage map in Figure 3.4.

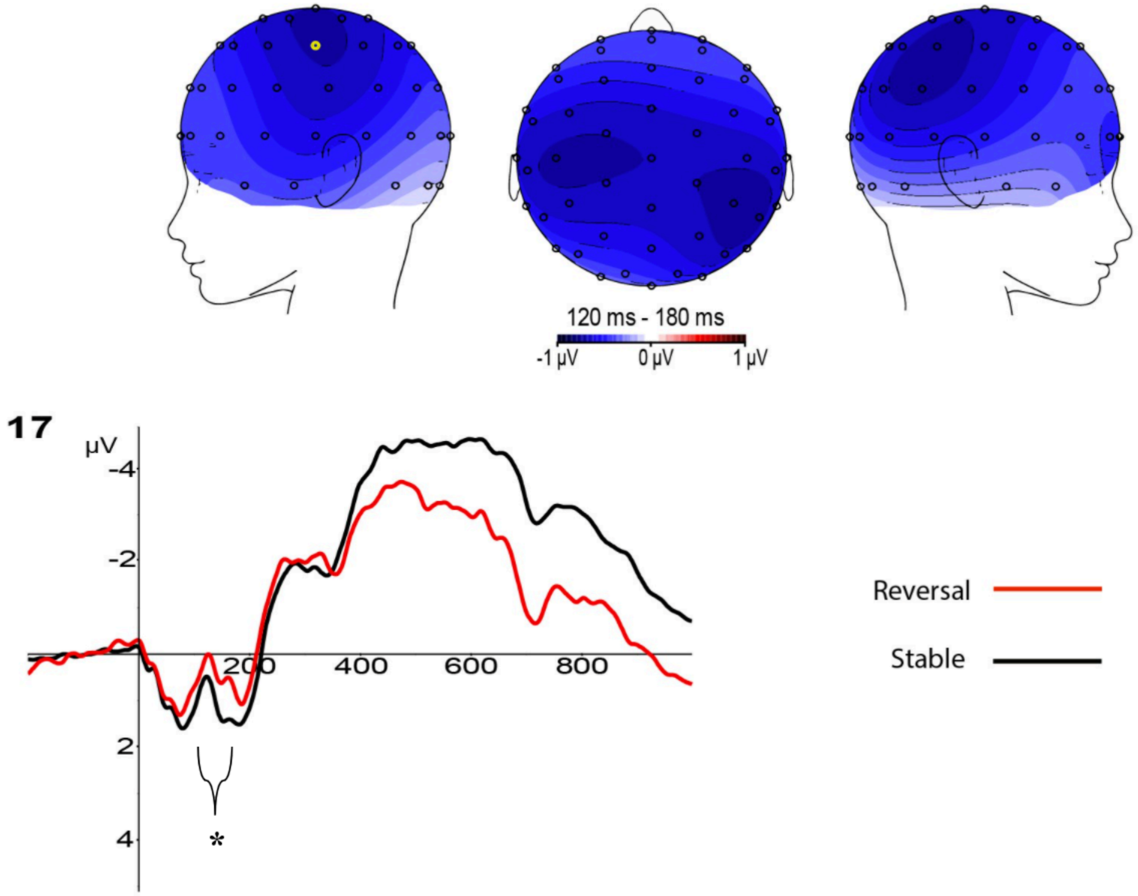


Figure 3.4 First Negativity. Voltage maps and electrode 17 grand averaged ERP. Voltage maps from a left, top, and right view depicting the mean amplitude difference between reversal and stable trials at the same time frame of 120ms - 180ms are shown on the top row. Electrode 17 (electrode C3 on the international standard 10-20 system) displayed the largest negative amplitude difference (μV) during this time window and is marked in yellow on the left-facing voltage map. An ERP of electrode 17 grand averaged across participants (N = 19) is shown in the bottom row. The time window at which the effect occurred is indicated by an asterisk.

3.2.3 Second Negativity

Mass univariate analysis of reversal minus stable trials (Figure 3.3) identified a second, longer lasting negativity evident from ~200 - 350 milliseconds post-stimulus at bilateral temporal locations, larger on the left hemisphere. It was determined that electrode 47 (electrode T7 on the international standard 10-20 system) detected the strongest negative amplitude difference (μV) during this time window, marked as yellow in the left-facing voltage map in Figure 3.5.

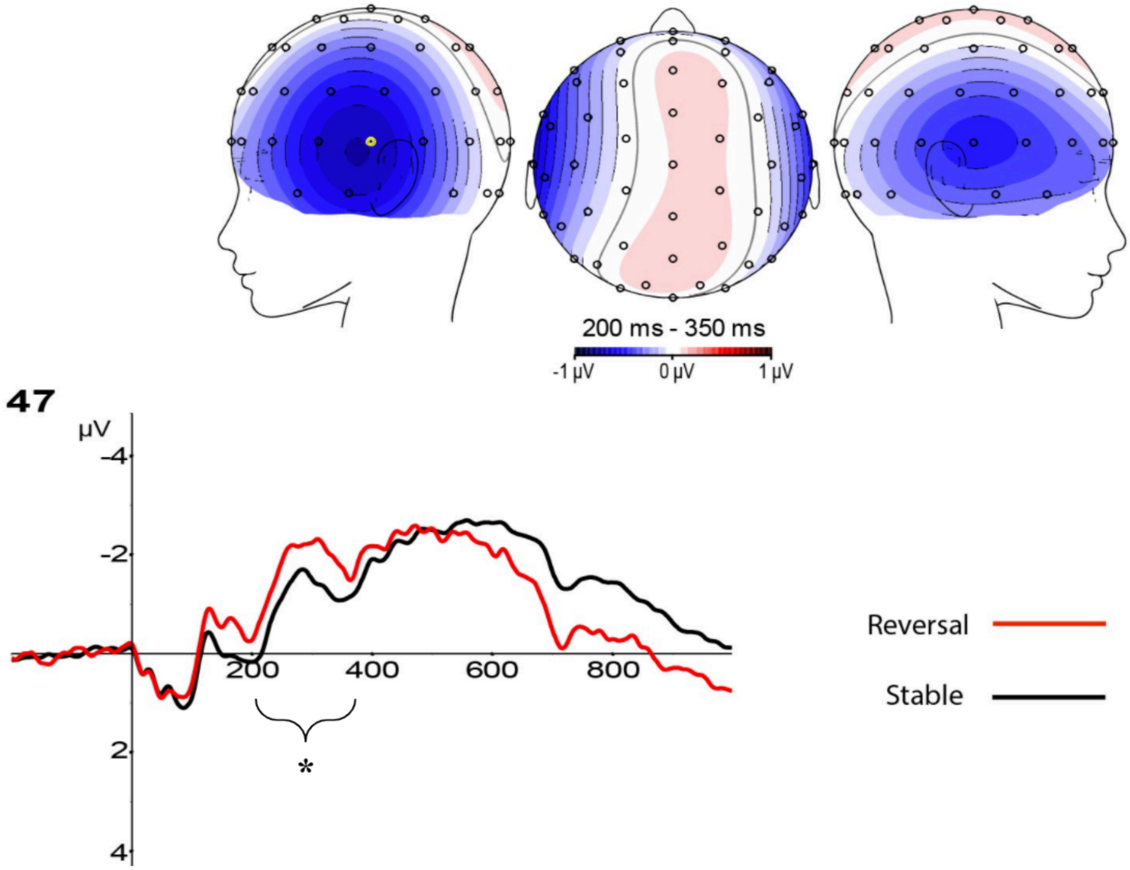


Figure 3.5 Second Negativity. Voltage maps and electrode 47 grand averaged ERP. Voltage maps from a left, top, and right view depicting the mean amplitude difference between reversal and stable trials at the same time frame of 200ms - 350ms are shown on the top row. Electrode 47 (electrode T7 on the international standard 10-20 system) displayed the largest negative amplitude difference (μV) during this time window and is marked in yellow on the left-facing voltage map. An ERP of electrode 47 grand averaged across participants is shown in the bottom row. The time window at which the effect occurred is indicated by an asterisk.

3.2.4 Late Positive Complex

Mass univariate analysis of reversal minus stable trials (Figure 3.3) identified a late positive complex evident from ~400 - 1000 milliseconds post-stimulus encompassing centro-parietal locations. It was determined that electrode 5 (electrode CPZ on the international standard 10-20 system) exemplified the strongest positive amplitude difference (μV) during this time window; marked in yellow on the top-view voltage map in Figure 3.6.

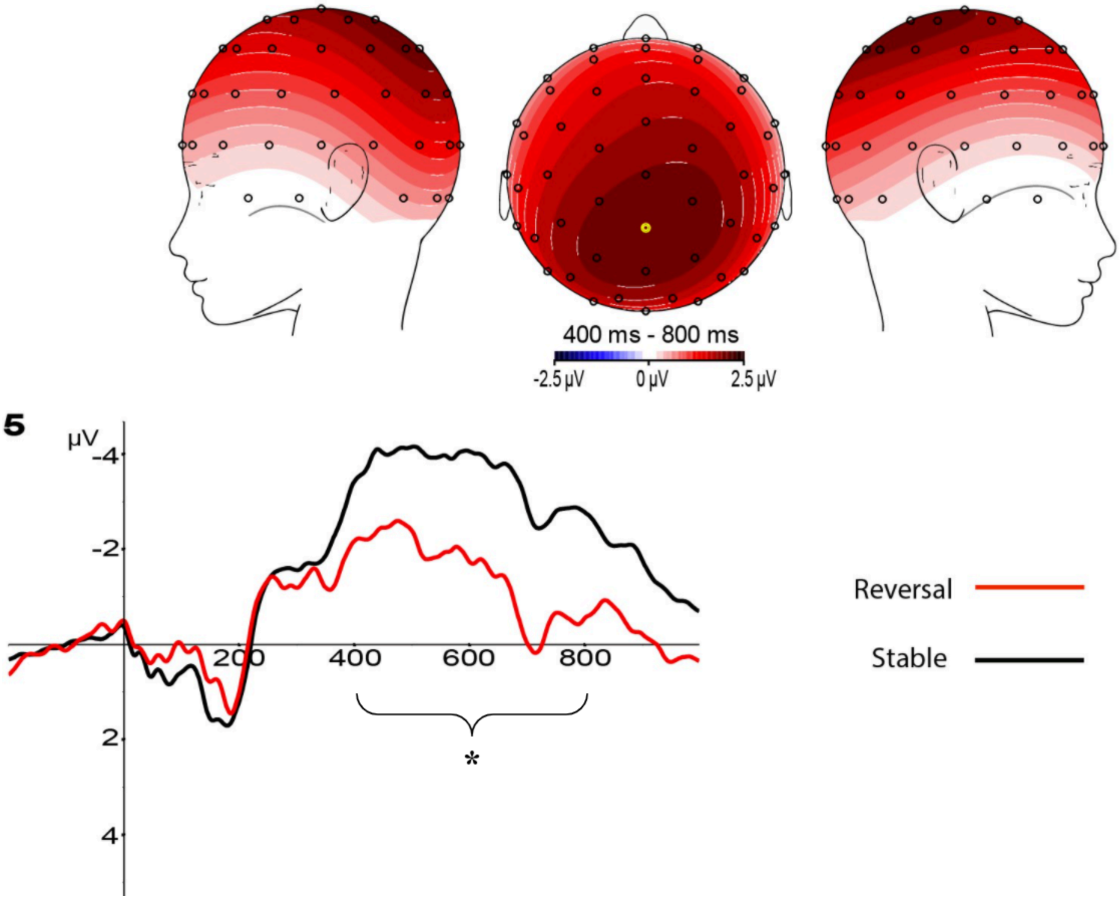


Figure 3.6 Late Positive Complex. Voltage maps and electrode 5 grand averaged ERP. Three voltage maps (left, top, and right views) depicting the mean amplitude difference between reversal and stable trials at the same time frame of 400ms - 800ms are shown on the top row. Electrode 5 (electrode CPZ on the international standard 10-20 system) displayed the largest positive amplitude difference (μV) during this time window and is marked in yellow on the top-view voltage map. An ERP of electrode 5 grand averaged across participants is shown in the bottom row. The time window at which the effect occurred is indicated by an asterisk.

3.2.5 Laurel vs. Yanny: Mass Univariate Analysis

No difference effects were detected in a false discovery rate mass univariate analysis of Laurel vs. Yanny ERPs at any electrode during any time-point from 0-1000ms.

3.3 Correlations between Behavioral & Electrophysiological Results

In order to evaluate potential correlations between behavioral and electrophysiological results for all three effects, mean amplitudes were individually calculated for each participant by averaging the amplitudes of the electrode cluster and across time-windows with the largest effect. Electrode clusters were determined by pinpointing the electrode with the largest amplitude difference and additionally selecting each of the surrounding electrode amplitudes. This calculated each participant's mean amplitude (μV) for each electrophysiological effect across the relevant time-windows (separately for the first negativity, second negativity, and late positive complex). Reversal rates were plotted in numerical order on the horizontal axis and mean amplitudes (μV) were plotted on the vertical axis to assess their correlation. One plot was generated for each of the three effects (Figures 3.8, 3.9, 3.10).

Statistical analysis in R-studio calculated R^2 and *p*-values for each linear model (as well as the slope and standard error) to assess whether reversal rate would predict the amplitude of each effect.

3.3.1 First Negativity: Mean Amplitude vs Reversal Rate Comparison

A statistically significant linear relationship between reversal rate and mean amplitude (μV) was detected in the first negativity, revealing that reversal rates did predict the amplitude of the effect ($R^2 = 0.168$; p-value = 0.04; slope = -2.07; Std Err = 0.96). Reversal rates were plotted in numerical order on the horizontal axis with mean amplitudes plotted on the vertical axis (depicted in Figure 3.7). The vertical axis was scaled from positive to negative due to the effect's negative amplitude (μV). Plots and statistical analysis were generated using R-studio.

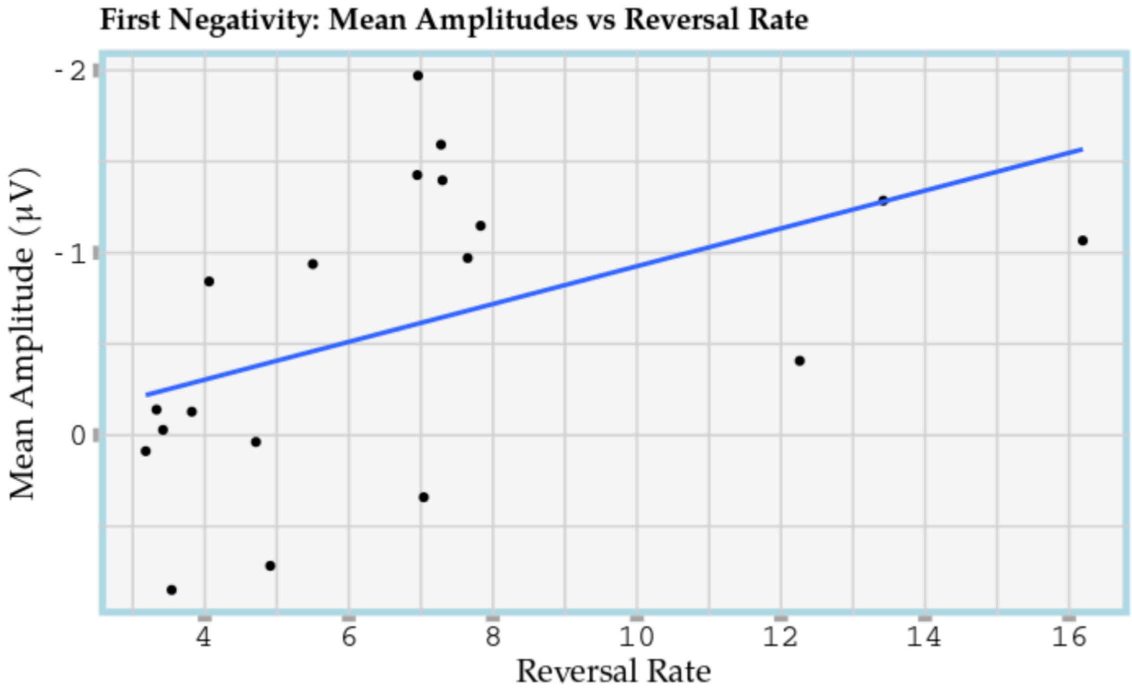


Figure 3.7 Mean amplitude (μV) versus reversal rate for the first negativity. A statistically significant linear relationship was detected, revealing that reversal rates did predict the amplitude of the negativity ($R^2 = 0.168$; p-value = 0.04 ; slope = -2.07 ; Std Err = 0.96). The vertical axis was scaled from positive to negative due to the effect's negative-going amplitude difference.

3.3.2 Second Negativity: Mean Amplitude vs Reversal Rate Comparison

The relationship between reversal rate and mean amplitude (μV) was not statistically significant in the second negativity, revealing that reversal rates did not predict the amplitude of the effect (R^2 = -0.054 ; p-value = 0.78 ; slope = -0.45 ; Std Err = 1.64). Reversal rates were plotted in numerical order on the horizontal axis with mean amplitudes plotted on the vertical axis (depicted in Figure 3.8). The vertical axis was scaled from positive to negative due to the effect's negative amplitude (μV). Plots and statistical analysis were generated using R-studio.

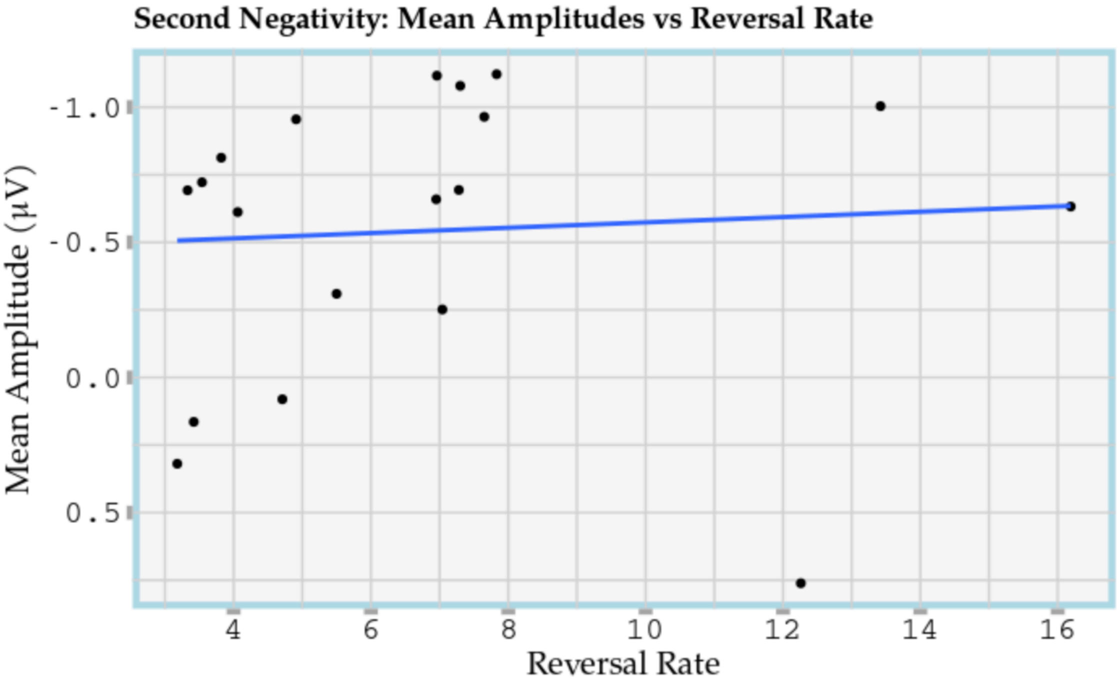


Figure 3.8 Mean amplitude (μV) versus reversal rate for the second negativity. The relationship between reversal rate and mean amplitude (μV) was not statistically significant in the second reversal negativity, revealing that reversal rates did not predict the amplitude of the effect (R^2 = -0.054 ; p-value = 0.78 ; slope = -0.45 ; Std Err = 1.64). The vertical axis was scaled from positive to negative due to the effect's negative-going amplitude difference.

3.3.3 Late Positive Complex: Mean Amplitude vs Reversal Rate Comparison

Regarding the late positive complex, the relationship between reversal rate and mean amplitude (μV) was not statistically significant; revealing that reversal rates did not predict the amplitude of the effect ($R^2 = 0.003$; *p*-value = 0.31 ; slope = -0.45 ; Std Err = 0.44). Reversal rates were plotted in numerical order on the horizontal axis with mean amplitudes plotted on the vertical axis (depicted in Figure 3.9). The vertical axis was scaled from negative to positive per-usual due to the effect's positive amplitude (μV). Plots and statistical analysis were generated using R-studio.

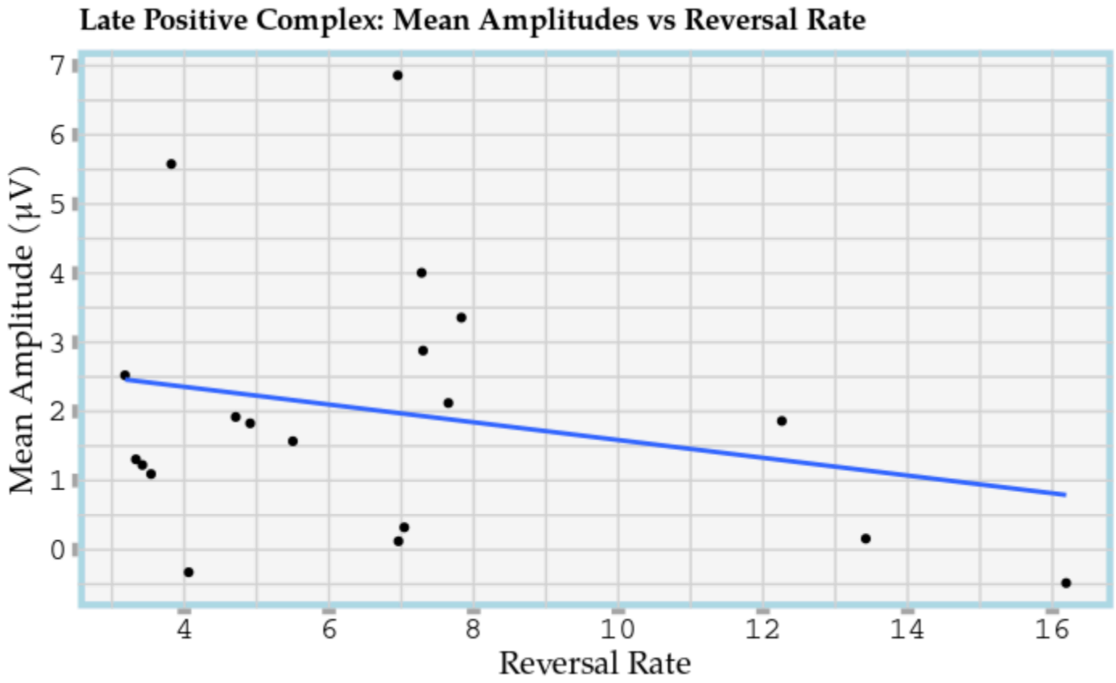


Figure 3.9 Mean amplitude (μV) versus reversal rate for the late positive complex. The relationship between reversal rate and mean amplitude (μV) was not statistically significant, revealing that reversal rates did not predict the amplitude of the effect ($R^2 = 0.003$; *p*-value = 0.31 ; slope = -0.45 ; Std Err = 0.44).

Chapter 4: Discussion

4.1 Summary of the Results

The present study sought to investigate electrophysiological components of auditory bistability associated with the perception of the bistable auditory language stimulus “Laurel & Yanny”; interpretable as either the name “Laurel” or pseudonym “Yanny”. The initial hypothesis was that an auditory reversal negativity (aRN) and an auditory late positive complex (aLPC) would be detected when participants experienced a perceptual reversal (i.e.; experienced a change in their conscious perception either from Laurel-to-Yanny or Yanny-to-Laurel) compared to perceptual stability (i.e.; experienced no change in perception, Laurel-to-Laurel or Yanny-to-Yanny). The results successfully support this initial hypothesis: two early negativities were found, followed by a late positive complex (aLPC) during perceptual reversals as compared to perceptual stability.

In an attempt to further interpret the identity of the three components found, the present study then questioned whether there would be any significant correlation between participant reversal rate and the amplitude (μV) of the three electrophysiological effects found. It was found that a statistically significant linear relationship was only detected for the first negativity, revealing that this component resembled an oddball effect similar to the mismatch negativity (MMN). Meanwhile for the second negativity and the aLPC, the relationship between reversal rate and mean amplitude (μV) was not statistically significant: reversal rates did not predict the amplitude of these effects. Thus, the second negativity (aRN) and the aLPC are more validly interpretable as neural correlates of changes in perception resulting from auditory bistability. In other words, these two components were present and approximately equal in magnitude regardless of how often a perceptual reversal was experienced.

A final question of the study sought to answer whether it would be possible to detect a neural difference between participants' reported perception of Laurel and their perception of Yanny. The results did not identify an electrophysiological difference between the perception of both percepts, at least within our initial analysis of averaged ERP amplitudes.

4.2 First Negativity: An MMN-like Effect

The first auditory negativity (aN) was evident from ~120 - 180 milliseconds post-stimulus, bilaterally distributed upon central and centro-parietal locations. Participants' reversal rates were plotted in numerical order and compared to the amplitude of the aN. A larger reversal rate reflects that it took longer to perceive a perceptual reversal, whereas a smaller reversal rate reflects that it took shorter to perceive a perceptual reversal. A statistically significant linear relationship was detected for the effect, revealing that as reversal rates increased (making reversals more rare); so did the amplitude (μV) of the negativity. These results are highly reminiscent of the mismatch negativity (MMN), which is an electrophysiological component predominantly generated by the auditory cortex (Alho, 1995). The MMN acquires an increasingly negative amplitude as the ratio between presentations of an oddball stimulus in a stream of repetitive stimuli increases (i.e.: the more infrequent the deviant stimulus amongst standard stimuli; the larger the MMN). This bears a strong resemblance to the aN, in that participants who perceived larger numbers of stable trials before a perceptual reversal had correspondingly larger aN amplitudes: the more infrequent the reversals, the larger the amplitude of the effect.

The MMN typically occurs when successive presentations of a stimulus are interrupted by an "oddball" stimulus; or a deviant form of the stimulus (Bartha-Doering et al., 2015). For example in the case of audition, if participants are repetitively presented with an identical auditory tone; disrupting successive presentations of the standard tone with an altered tone (e.g.; the same tone with a different pitch, volume, or timbre) will elicit an MMN (Näätänen et al., 1989). This is of interest, due to the fact that there was only one auditory stimulus played repeatedly in all blocks of the present study, and no oddball stimuli or altered versions of the stimulus were used during the experiment. A 2015 review paper by Bartha-Doering et al. stated that “the MMN requires a stimulus deviance to occur, whereas it cannot be elicited by a single stimulus, the first stimulus in a series, or by a deviant stimulus alone” (introduction par. 6). Although this may be the case for many auditory stimuli, the present study offers up the notion that the dually perceivable nature of certain auditory bistable stimuli may also elicit an MMN-like effect as the timing and scalp-distribution of the aN appear to be very similar to that of the MMN: the MMN component is often detected between 100 and 200 milliseconds with a frontal to central scalp distribution (Shtyrov & Pulvermüller, 2007 ; Bartha-Doering et al., 2015). Because reversal trials were perceived less so than stable trials in all participants (“Reversal” trial mean = 27% compared to “Stable” trial mean = 73%), they constitute as the only possible form of a “stimulus deviance”; although as all perceptual reversals occurred endogenously this deviance was solely a perceptual phenomenon.

To date, instances of the MMN in EEG studies of auditory bistability have seemingly not yet been documented (Davidson & Pitts et al., 2014 ; Higgins et al., 2020). Davidson & Pitts (2014) located an aRN when participants reported experiencing perceptual reversals while listening to auditory bistable streaming stimuli. A similar behavioral analysis was performed: the size of participant’s reversal rates were compared to the amplitude of their aRN. Having found no effect, it was suggested that the aRN and MMN are distinct components (Davidson & Pitts et al., 2014). The results of the present study seem to validate this hypothesis, as the early MMN-like aN occurred independently, subsequently followed by a secondary negativity that showed no correlation with participant’s reversal-rates (*p*-value = 0.78); this second component thus likely being an aRN (auditory reversal negativity). Both negative effects were detected at different spatial and temporal locations; the aRN being evident from ~200 - 350 milliseconds post-stimulus onset, bilaterally distributed upon the temporal lobe with a larger amplitude detected in the left hemisphere. If other auditory bistable stimuli have the potential to elicit an MMN-like effect, these results seem to suggest that the MMN would occur independent of an aRN component; perhaps depending upon the manner or extent of a bistable stimuli’s perceptual deviance from its priorly perceived state.

The present study evidently being the first to investigate electrophysiological components of auditory language bistability, it may be entirely possible to observe the MMN (followed by an aRN) due to the nature of how lexical phonemes are perceived. The scalp distribution of the MMN is found to be bilateral in frontal or central regions, although it is often found to favor a larger amplitude in the right hemisphere (Paavilainen et al., 1991). However, when the MMN results from a stimulus that bears a phonetic deviance from prior repetitions of another auditory stimulus, the MMN response amplitude is instead greater in the left hemisphere; a finding that has been theorized to occur due to the existence of language processing structures found in left hemispheric regions (Näätänen et al., 1997; Shtyrov et al., 2000 ; Zevin et al., 2010). In the present study, the MMN-like effect was found centrally and bilaterally, however the amplitude of the response was found to be greater in the left hemisphere at electrode C3 on the international standard 10-20 system. This evidence suggests that the aN component was a response to perceived phonetic changes in the stimulus of the present study.

Studies have shown that categorical perceptions of Chinese lexical tones have previously been revealed by the MMN (Xi et al., 2010). Furthermore, a review paper written by Shtyrov & Pulvermüller published in 2007 interpreted the relationship between language and the MMN, stating that due to a "growing body of evidence, the pattern of MMN response to individual word deviants has been considered as indicating a word's 'signature,' it's memory trace activated in the brain" (Shtyrov & Pulvermüller, 2007, lexical processes, par. 3). This theory would support the notion that the MMN-like effect found in the present study is responsible for playing a role in the indication of early phonetic differences between "Laurel" and "Yanny"; the early timing of the component suggesting a link to the detection of perceptual deviance resulting from lexical processing differences in the first phoneme of each name: <u>La</u>urel and <u>Ya</u>nny.

An increase in the response amplitude of the MMN has been identified when individual words are the deviant stimulus, as compared to when pseudonyms are the deviant stimulus (Shtyrov et al., 2005 ; Shtyrov & Pulvermüller, 2007). Similarly, words that are more meaningful exhibit a greater MMN response amplitude as compared to meaningless pseudonyms; theorized to occur due to the stimulation of cortical memory traces linked to more familiar words as a result of more robustly connected neural populations (Bartha-Doering et al. 2015 ; Pulvermüller & Shtyrov, 2006 ; Shtyrov & Pulvermüller, 2007). An interesting analysis to perform with the data from the present experiment would be to isolate the response amplitude of the aN during perceptual reversals that occurred when participants switched from "Yanny-to-Laurel", and compare this with the response amplitude of the aN during perceptual reversals that occurred when participants switched from "Laurel-to-Yanny". Given that "Laurel" is a more common name and word in the English language, perhaps an increased MMN response amplitude would be detected when participants experienced a perceptual reversal to Laurel (from Yanny) rather than a perceptual reversal to Yanny (from Laurel). It may also be the case that no differences in the aN effect would be detected, say for example if participant's expectations of "Yanny" as an anticipated interpretation of the stimulus potentially could render it as meaningful and familiar as "Laurel" during the experiment.

An MMN-like effect resembling the aN from the current study presumably has not yet been located due to the lack of previous experimentation on auditory language bistability. Exogenously induced perceptual reversals of a bistable Necker cube have been argued to elicit the MMN (Urakawa et al., 2018), however this shows that the effect relied on the facilitation of perceptual alterations and thus some form of a stimulus deviance. As discussed, the ambiguity that "Laurel & Yanny'' delivers to participants involves perceptual processing differences related to language phonemes. Having established a strong correlation between the MMN's relationship with deviances in lexical and phonological qualities, the present study's use of an auditory bistable language stimulus seemingly isolated a novel auditory processing sequence that implements all three effects; the MMN-like aN, the aRN, and the aLPC.

4.3 Second Negativity: Auditory Reversal Negativity

A second, longer lasting negativity was evident from ~200 - 350 milliseconds post-stimulus at bilateral temporal locations. Unlike the first negativity located in the present study (aN), reversal rates did not predict the amplitude of this second component; and the effect was also distinct from the MMN-like aN both spatially and temporally. This component was followed by a late positive complex (LPC), suggesting that this second negativity is an auditory reversal negativity (aRN); an auditory analog of the visual reversal negativity (RN). The RN has been reported in many studies during perceptual reversals resulting from visual bistable stimuli (Abdallah & Brooks, 2020 ; Britz et al., 2009 ; Britz & Pitts, 2011 ; Intaite et al., 2010 ; Kornmeier & Bach, 2004, 2005; Kornmeier et al., 2007; Pitts, 2007, 2009). It bears a similar timing to the aRN, occurring between 150 - 350 milliseconds post stimulus onset (Davidson & Pitts, 2014 ; Pitts et al., 2008). The RN appears independent of task and response style, and is found upon scalp distributions relative to regions that process the related modality of the bistable stimulus that they are elicited from (Abdallah & Brooks, 2020). The finding of the aRN is a replication of results from previous auditory and visual experiments studying the EEG neural correlates of bistability, thus being in support of the initial hypothesis.

4.3.1 Previous Accounts of the aRN

Neural correlates pertaining to auditory bistability are considerably under explored in EEG studies. Two previous accounts of an auditory reversal negativity (aRN) have been documented in EEG studies of auditory bistability (Davidson & Pitts, 2014 ; Higgins et al., 2020); both instances were found during perceptual reversals that participants experienced while listening to bistable auditory streaming stimuli.

Auditory streaming stimuli differ from “Laurel & Yanny” in how they are perceptually bistable, as they are not spoken language stimuli. Both studies used distinct forms of auditory streaming stimuli which differed characteristically in the perceptual sensations they elicited in participants; both of which were constructed from numerous tones (Davidson & Pitts, 2014 ; Higgins et al., 2020). The streaming stimuli presented in Davidson & Pitt’s study (2014) elicited the perceptual sensation of either ascending or descending pitch motion; and were complex Shepard tones built from six different sinusoidal harmonics. The streaming stimuli presented in Higgins et al. (2020) were triplets of ABA tones followed by a brief tone absence, dually perceivable as either one stream eliciting an interleaved "galloping" sonic sensation; or as two separate streams eliciting a repetitively metronomic tone.

In both studies, an aRN was found during perceptual reversals distributed upon frontal to fronto-central scalp locations (Davidson & Pitts, 2014 ; Higgins et al., 2020). The location of the aRN in the present study differed from these; appearing with a bilateral distribution upon temporal lobe locations. Upon further examination of the present study's aRN, it was found that the response amplitude was greater in the left hemisphere rather than on the right; centered upon electrode T7 (on the international standard 10-20 system). This reveals a key topographic difference in the processing of perceptual reversals resulting from "Laurel & Yanny" as compared to those resulting from auditory streaming stimuli; plausibly suggesting the recruitment of language processing structures housed in the left hemisphere (Shtyrov et al., 2000 ; Zevin et al., 2010). This follows in the results of prior research, exemplifying that the scalp distribution of the visual RN and now the aRN both appear to selectively inhabit topographies related to regions that process the sensory modality of the stimuli they are elicited from.

4.3.2 The Role of the aRN: Interpretations Based on the Visual RN

Because the aRN in the present study appeared during reversal trials as compared to stable trials, it can be theorized to be a neural marker indexing the transition from one perceptual state of the stimulus (e.g.; Yanny) to the other perceptual state (e.g.; Laurel). How the aRN is involved in transitions between perceptual reversals is up for debate. As the aRN possesses many similarities to the RN, it is possible to theorize the underlying mechanisms of this auditory component based on its visual predecessor.

Attempts to parse the role of the RN have led to a discussion regarding whether the component is predominately implicated in awareness, selective attention, or either conscious or unconscious perception (Intaite et al., 2010). Greater discussions challenge the notion that these concepts are separate entities at all (Davidson & Pitts, 2014 ; Tannenbaum, 2001 ; Tsuchiya & van Boxtel, 2013).

Information processing theories (i.e.; bottom-up and top-down processing) offer angles by which to interpret the mechanisms behind the RN's link to changes in perception. It has long been debated how bistable perception may be mediated via bottom-up processing or top-down processing, two processes that may not be entirely mutually exclusive (Abdallah & Brooks, 2020 ; Intaite et al., 2010 ; Pitts et al., 2008). For example, if perceptual reversals are solely brought about by bottom-up processes; an early adaptation (decrease in response) of sensory neural populations would explain switches in perception (Kornmeier & Bach, 2006 ; Toppino & Long, 1987). In the case of top-down processing, this adaptation would take place in higher order processing regions of the brain; although as discussed by Kornmeier & Bach (2006), "...few ERP studies on ambiguous perception mirror the situation with psychophysical results in either supporting the top-down approach or the bottom-up approach" (introduction, par. 3).

How the aRN may relate to top-down processing can be inferred from studies that have explored bistability and visual perceptual reversals under the manipulation of participant's voluntary top-down control. An fMRI study published in 2005 by Slotnick & Yantis discovered that both voluntarily induced perceptual reversals as well as the voluntary mediation of attention to different regions of a visually bistable Necker cube were accompanied by increased activity in the posterior parietal cortex; implying that voluntarily induced perceptual reversals (top-down control) are mediated by spatial attention. It has additionally been found the RN increases in amplitude when participants volitionally change their perception from one percept to the other, as compared to when participants simply passively perceive a bistable stimulus (Pitts et al., 2008). These findings imply that the RN plays a role in selective attention, as voluntary top-down processes are capable of strengthening the effect of the RN and make use of regions that are similarly recruited by selective attention (for a more detailed discussion on the matter, refer to Pitts et al., 2008).

Although this data suggests the RN may be implicated in attention somehow, Intaite et al. (2010) found that the RN is not an analog of the N2pc; another common component known to index visual selective attention (Hickey et al., 2006). Yet, as there still appears to be an influence of top-down control over the RN; this suggests a need for more research on the matter. As remarked by

Davidson & Pitts (2014), in order to gain more certainty of the RN as a non-attentional correlate; attentional shifts that precede the RN and perceptual reversals must be studied in greater detail; as preceding activity in the right inferior parietal cortex has been found to be associated with perceptual reversals of a visually bistable Necker cube (Britz & Michel, 2009).

It is not too surprising that the RN is neither an analog of the N2pc, nor does it seem to resemble this attentional component (Intaite et al., 2010). This is because perceptual reversals are not always under a participant's top-down control, and will occur involuntarily; even inevitably so (e.g.; bottom-up processing). In this case, it can be suggested that attention is modulated by other processes; perhaps by unconsciously driven perceptual processes that may later give rise to conscious perception. Thus, perhaps the RN is a component responsible for processing changes in the perceptual structure or general composition of a bistable stimulus (Abdallah & Brooks, 2020).

If it is true that the RN is a marker indicative of *unconscious* perceptual processing, it must also be the case that it is not an analog of ERP components proposed to index conscious awareness, like the visual awareness negativity (VAN). Intaite et al (2010) suggested that if the RN was a subset of the VAN, it would similarly be detected in response to the characteristics of a stimulus that enter visual awareness. No correlation was found between the RN and the content of participant's visual awareness however, suggesting that the RN is an unlikely candidate for awareness (Intaite et al., 2010). Further studies should examine this theory more closely to determine whether the RN may somehow be related to conscious perceptual processing.

Exactly how the RN relates to the VAN and the LPC (late positive complex) is of great interest, as multiple theories exist; many of which attempt to argue how the three may be implicated in perceptual processing, or phenomenal consciousness and access consciousness; to name a few (Eklund & Wiens, 2018). This is also of interest because both the VAN and the RN are often followed by the LPC in many different experimental paradigms (Koivisto & Revonsuo, 2010 ; Eklund & Wiens, 2018). More deliberation on the relationship between the aN, the aRN, and the aLPC of the present study is available in section "4.5 Interpreting Results via the VAN & Theories of Consciousness".

The latency of the RN is sensitive to an experiment's response interval length, as are the scalp distribution and the timing of the effect (Abdallah & Brooks, 2020 ; Kornmeier & Bach, 2012). This further suggests that the RN can be modulated by experimental idiosyncrasies and is not solely modulated by perception (Abdallah & Brooks, 2020).

Interestingly, certain studies have reported that the RN not only occurs during endogenous perceptual reversals; it has also been documented during exogenous (stimulus driven) perceptual reversals (Kornmeier & Bach, 2006). This brings into question whether studies implementing exogenously produced reversals may be confusing the MMN for the RN (or vice versa), as another study claimed that exogenously induced perceptual reversals of a bistable Necker cube elicited the MMN (Urakawa et al., 2018). As discussed, differences in the MMN and the RN can be detected by analyzing timing of the effect, scalp distribution, and whether reversal rates predict the amplitude of the effects in question.

Given what is known of the RN and its similarity to the aRN, it seems reasonable to postulate that the aRN noted in the current study is a correlate indicative of unconscious perceptual processing, coinciding with involuntarily experienced perceptual reversals of "Laurel & Yanny".

4.4 The Auditory Late Positive Complex

The auditory late positive complex (aLPC) was evident from ~400 - 1000 milliseconds post-stimulus onset, encompassing centro-parietal locations. This finding is in support of the initial hypothesis and successfully replicates the results of other EEG paradigms that have investigated the neural correlates of bistable stimuli. It was determined that electrode CPZ on the international standard 10-20 system exemplified the strongest positive amplitude difference (μV) during this time window. Previously, only one other instance of an aLPC located during perceptual reversals resulting from bistable auditory stimuli has been documented by Davidson & Pitts (2014). The aLPC of the present study was much longer lasting than that of Davidson & Pitts, appearing from ~400 - 1000 milliseconds as compared to ~320 to 380 milliseconds in length. Both components were found centered on the midline, however their locations slightly differed from one another; with Davidson & Pitt's aLPC appearing fronto-centrally while the current aLPC appeared centro-parietally (Davidson & Pitts, 2014).

In previous EEG studies exploring visual bistability, the late positive complex (LPC) has routinely been found ~350 milliseconds post stimulus onset at centro-parietal regions; subsequent to the RN (Kornmeier and Bach, 2004; Pitts et al., 2007; Koivisto and Revonsuo, 2010). As previously discussed, the RN appears independent of task and response style during voluntary and involuntary perceptual reversals (Abdallah & Brooks, 2020). This is not the case for the LPC however, which is instead task-dependent; appearing when participants are asked to report or recall their perception of a stimulus (Yang et al., 2019). Thus, it has been proposed that the LPC may be implicated in awareness, memory, and reporting perception (Wilenius & Revonsuo, 2007 ; Salti et al., 2012 ; Eklund & Wiens, 2018).

4.5 Interpreting Results via the VAN & Theories of Conciousness

The LPC does not only appear exclusively after the RN. As previously mentioned, there exists a component called the VAN (visual awareness negativity) which bears strong resemblance to an enhanced N200 component and has been theorized to be an early neural correlate of awareness (Eklund & Wiens, 2018). The amplitude of the VAN becomes increasingly more negative correlative to the features and magnitude of stimuli that participants report as aware versus unaware (Eklund & Wiens, 2018 ; Koivisto & Revonsuo, 2010 ; Ojanen et al., 2003). The VAN is also often followed by the task-dependent LPC; which resembles an excitatory analog of the P3 component (Eklund & Wiens, 2018) and also increases in positivity correlative to stimuli reported as aware versus unaware (Wilenius & Revonsuo, 2007). This evidence has also been noted in proposed auditory analogs of the visual VAN and LPC; as one recent study found an auditory awareness negativity (AAN) and subsequent aLPC that also correlated with stimulus awareness in similar ways (Eklund & Wiens, 2019).

Researchers are attempting to explain the role of these components given the current data through several theories of consciousness. A group of philosophies both formerly and recently advocated in current literature are proposed under the name “global workspace theory” (Baars, 2005 ; Baars et al., 2021 ; Dehaene and Changeux, 2011 ; Mashour et al., 2020), which postulates that sensory information is processed unconsciously amongst respective pathways, until it arrives at a neural network located at frontal and parietal cortical areas (Higgins et al., 2020). This network is densely composed of long-range axons which subsequently make this information globally available to a multitude of brain regions for conscious access (Dehaene and Changeux, 2011). Through the lens of global neuronal workspace theory, the VAN has been theorized to represent unconscious processing; followed by an LPC component indicative of an update to the global workspace and thus signifying the LPC as a marker of visual awareness (Eklund & Wiens, 2018 ; Dehaene and Changeux, 2011).

However, the results of a 2018 study by Eklund & Wiens support the notion that the VAN does not index unconscious processing, providing evidence that argues the component is an early correlate of awareness. Koivisto et al (2016) also provide evidence suggesting that this is the role of the VAN, and postulate that the LPC does not play a role in visual awareness; perhaps rather indexing post-perceptual processing of stimuli or other another mechanism. This evidence may suggest that unconscious processing might take place at a different timing than that of the VAN (perhaps previously or in some concurrent form), or that the VAN inhabits a more local range of awareness while the LPC inhabits a more global range.

This latter idea may fall into another theory of consciousness known as the “recurrent processing theory” (Lamme & Roelfsema, 2000), which postulates that sensory cortical processing gives rise to “phenomenal consciousness” (the present test conscious experience of a given moment), while the processing of information between frontal parietal cortices and the sensory cortex gives rise to “access consciousness” (globally available consciousness, e.g., the use of memory in reasoning) (Eklund & Wiens, 2018, introduction par. 4). In this case, the VAN may reflect phenomenal consciousness while the LPC may reflect access consciousness; although Eklund & Wiens (2018) and previously mentioned studies have not provided such evidence or tested whether this may be the case.

In order to refute global neuronal workspace theory, studies would have to find evidence of a VAN coinciding with awareness; but not followed by an aLPC.

Because the RN has not yet been found to correlate with awareness in the same manner as the VAN (Intaite et al., 2010), it may be inferable that the RN and aRN are both correlates indicative of perceptual processing, more specifically unconscious perceptual processing. Generally speaking, more evidence is needed to establish a relationship between the VAN and the RN, as there are slight similarities between the two (e.g.; their timing and relation to a subsequent LPC) (Eklund & Wiens, 2018). If it is the case that the VAN is an early neural correlate of awareness as new evidence seems to suggest; it would be very enticing to attempt to determine if the RN and aRN may be more related to awareness than we might think. If evidence was found to support the notion that the RN was somehow implicated in early awareness as well, then the LPC may reflect either access consciousness or it may be indicative of post-perceptual processing. Overall, the question remains as to what role the RN plays in attention, awareness, and perceptual processing.

Applying global neuronal workspace theory to all three auditory components found in the present study, the first negativity likely represents an MMN-like effect responding to early sensory detection of a perceived phonetic deviation in the stimulus. The subsequent aRN may then index some form of unconscious perceptual processing of “Laurel & Yanny”; perhaps integrating more lexical and auditory configurations of the stimulus. The following aLPC effect may then reflect an update to global workspace, thereby giving rise to auditory and language awareness and resulting in the ability for a participant to respond based on their given perception of the stimulus: “Laurel” or “Yanny”.

Applying this model to experimental trials of the present study, the most recent aLPC (i.e.; global update) would be responsible for the generation of a new motor system plan in order for the participant to report a perceptual reversal. Say for example that the participant had heard and reported numerous stable trials of “Yanny” previously, but then experienced a sudden perceptual reversal trial to “Laurel”. In this case, the aLPC would theoretically inform the inhibition of the participant’s prior motor response (i.e.; their finger pressing the

"Yanny" button); influencing the decision to now press the "Laurel" button instead.

The stimulus is ~670 milliseconds long, thus if the aLPC effect does reflect an update to the global workspace giving rise to auditory awareness, this theory would suggest that this change in awareness would occur at the earliest ~400 milliseconds into listening to the stimulus; by which point one stimulus iteration would be over halfway through being listened to. Given that it was not out of the ordinary for participants to quickly respond as the stimulus began (seemingly before the stimulus was half-over), at first glance it may seem intuitively surprising that conscious awareness of the percept would occur only after the stimulus was halfway through being played. This is one of the hallmarks of global neuronal workspace theory; which postulates that conscious awareness is delayed by a range of ~300 – 400 milliseconds (Baars, 2005 ; Baars et al., 2021). Some disagree with the theory that the LPC gives rise to awareness, instead proposing that it indexes confident reporting of the stimulus (Wilenius & Revonsuo, 2007). On the other hand, a few studies have unabashedly posited that the P3 component (in this context the aLPC) reflects conscious processing of a stimulus, regardless of its occasional relatively late timing (Lamy et al., 2009 ; Salti et al., 2012). Salti et al (2012) suggestively refute claims that the P3 component might relate more so to "confident" reporting of the stimulus; providing some evidence on the matter.

In the case of the current study, it is important to deliberate on just how much of the stimulus can be heard before 400 milliseconds occurs, or rather even before 300 milliseconds. Given that the first syllable of "Laurel & Yanny" takes up a larger duration of time as compared to the second, it does not seem unreasonable for it to be the case that conscious awareness of the percept might occur beginning at 400 milliseconds post stimulus onset; yet it remains to be a subject of interest as current theories seem to disagree on the role and meaning of the LPC. Further examination into the role of the LPC and the aLPC should be done in order to determine whether or not the component may be related to post-perceptual processing, reporting of a stimulus, access consciousness, or an update to the global workspace giving rise to conscious awareness.

Attempting to apply the recurrent processing theory to the present study's results instead of global neuronal workspace theory comes with its challenges. This is simply because as previously discussed; there is little evidence providing an argument for the RN as a neural correlate of early awareness (Intaite et al., 2010); thus rendering any attempt to label it as phenomenal consciousness one without any prior evidence. Some have proposed that data from Koivisto et al. (2016) seemingly suggest that the LPC is possibly a non-aware component or a post-perceptual one; negating evidence from other studies and adding leverage to the current debate on theories of consciousness as well (Eklund & Wiens, 2018 ; Koivisto et al., 2016).

4.6 Laurel vs. Yanny Analysis

Another question presented by the current study sought to answer whether it would be possible to detect a neural difference between participants' reported perception of Laurel and their perception of Yanny. The results did not identify any electrophysiological differences between the perception of both percepts. These null results were identified via a false discovery rate mass univariate analysis, which was the same test used to detect differences in the reversal minus stable comparison of the experiment. It may seem quite odd that no differences were found, as there absolutely must be a difference in the perception of "Laurel" as compared to the perception of "Yanny" in the brain. Thus, these results do not reveal that there is no difference neurally; rather they suggest that the analysis used to identify a difference was not sensitive enough to detect one. Other forms of analysis exist that are much more sensitive to brain wave idiosyncrasies delineating the perception between two stimuli, such as intracranial EEG or multivariate pattern classification analysis; otherwise known as "decoding".

Decoding would likely be more capable of determining how participants' perceptions of Laurel vs. Yanny are distinguishable from the spatial distribution of scalp electric fields at each time-point (i.e.; time-resolved EEG decoding). For this analysis, pattern classifiers can be trained using linear support vector machines on a larger portion of the data, labeled as Laurel vs. Yanny. In other words, 80% of the Laurel vs. Yanny EEG data could be analyzed.

Following this analysis, the next step would then be to test the accuracy of the trained classifier on distinguishing what was perceived on individual trials by inputting the remaining dataset into the system. Put differently, the remaining 20% of the Laurel vs. Yanny EEG data could be run through the trained analysis to test how accurately the analysis could distinguish between the 20% based on the priorly analyzed 80%.

Other EEG studies implementing the use of auditory bistable stimuli have successfully uncovered neural differences between the perception of each percept via decoding (Billig et al., 2018). It would be very worthwhile to perform this more sensitive analysis on the Laurel vs. Yanny EEG data in order to more formally attempt to uncover neural processing differences between each name.

4.7 Discussing Behavioral Results

It has been theorized that as long as both stable states of a bistable stimulus are equivalently likely to take place; one will inevitably experience a perceptual reversal due to perceptual competition (Kornmeier & Bach, 2005). This theory was validated upon examining behavioral perception of both stimuli, as both “Laurel” and “Yanny” were equally perceived: “Laurel” was perceived an average of 49.82% of the time across all trials, while “Yanny” was perceived an average of 50.17% of the time. These percentages are a mean of all trials; including those that were present pre-artifact rejection.

In regards to perceptual reversals, “Stable” trials were perceived 3.17 times more than “Reversal” trials: “Stable” trials were perceived 73% of the time while “Reversal” trials were perceived 27% of the time. This was reflected in the study by reversal rates, which initially showed that it was taking a handful of participants too long to switch in perception (N = 4).

Pretest screening processes were implemented to vet individuals and select those who had faster reversal rates for experimental eligibility. These measures were highly successful at enrolling participants who yielded larger amounts of reversal trials, as only one other participant failed to reach the minimum threshold of ~70 reversal trials necessary for inclusion in a given condition. Those participants that were partially excluded solely due to a lesser than desired number of reversal trials (N = 2) were simply used for the Laurel - Yanny analysis. For more specific information regarding participants who were included and excluded from the study, refer to section "3.1.2 Participant Exclusion" of the Results.

4.8 Pretest Data

24 participants passed the pretest and made it through the entirety of the experiment, five of which were excluded from the study. Of the initial 24 participants, pretest data revealed that stimulus variant 8 was found to be the most bistable stimulus in the study; which was tuned -3 semitones lower than that of the original stimulus. Interestingly, the original unedited stimulus (variant 8) was only used in the study once.

Preliminary pilot results revealed that tuning the original stimulus to lower semitone frequencies more often influenced the perception of "Yanny", whereas tuning to higher frequencies more often influenced the perception of "Laurel". No other auditory adjustments were made to the stimulus apart from their initial semitone tuning, and yet these minimal frequency changes greatly influenced the listener's perception of the stimulus.

For the most part, stimulus variants chosen by participants held a roughly central distribution (see Figure 3.1 of the Results); although there was no symmetry to the distribution as there were several outliers and a decent amount of variance. Although 24 participants passed the pretest and made it through the entirety of the experiment; only 19 participants were included in all analyzed conditions of the EEG analysis. This would minimize the variance to a slight extent, as the majority of participants whose EEG data were used in the analysis selected and used centrally distributed stimulus variants during the experiment.

Because there was so much variation in the stimuli chosen during the pretest and subsequently used during the experiment, it would be interesting to perform an analysis comparing participant's stimulus frequency to the reversal minus stable analysis of their EEG data in an effort to examine whether there was any correlation. For example, if certain frequencies were found to generate a larger effect amplitude or a specific scalp topography, this would have interesting implications for the priorly discussed electrophysiological components (the MMN-like aN, aRN, and aLPC).

Figure A3 of the Appendix details the 19 participants who were included in all analyzed conditions of the EEG experiment and their respective semitone frequencies used. Additional information on stimulus frequencies can be found in the "3.1.1 Pretest Results" section of the Results.

4.9 Limitations

During actual experimental blocks, one recurrent and bothersome behavioral limitation was reported by a few participants. Participants were routinely checked on in between blocks and were asked how the experiment was going for them; as well as what they were perceiving. 5 participants whose EEG data were used in the final analysis reported that roughly halfway through the experiment, they were not hearing "Laurel" or "Yanny" and rather heard the first syllable of one name followed by the second syllable of the other; for example "Yan-el" or "Laur-y".

This was an alarming finding, given that perceiving a combinatory version of both percepts lowered the validity of using "Laurel & Yanny" to research neural correlates of bistable perception; possibly perturbing the experimental goal. When this occurred during a block, these participants were instructed to take a longer break and were offered water, candy, or tea. They were also instructed to report their initial perceptual instinct of the stimulus, and to not respond if uncertain. Each of these participants (N = 5) found that taking longer breaks in between blocks decreased the rate at which they heard combinatory versions of both percepts.

As mentioned in previous sections, another limitation was the initially low rate of perceptual reversals that certain participants had in the beginning of data collection. As discussed, this limitation was easily amended by configuring the pretest to also be an eligibility test for participation in the EEG experiment.

4.10 Future Research

There are many questions that could be explored in future experiments. However, before advancing onward to new avenues of research; it would be critically important to replicate the results of the current study as well as those of other novel studies exploring auditory bistability. EEG neural correlates of auditory bistability are underexplored compared to those which have studied visual bistability (Davidson & Pitts et al., 2014). Because of this, it can be somewhat discouraging to attempt to reflect all that the aRN and aLPC may represent to each other through their visual counterparts (the RN and LPC); as is the case for other more recently discovered auditory ERPs (i.e.; the AAN and N2ac). Ideally, future experimentation could firmly support current notions that these auditory components are analogs of their hypothesized visual counterparts. In the case of the aRN, it would be necessary to provide evidence that the aRN is unrelated to the N2ac (an auditory analog of the N2pc named the anterior contralateral component ; Gamble & Luck, 2011) in the same manner that the RN is not related to the N2pc (Intaite et al., 2010). Future experiments could also examine whether the amplitude of the aRN can be increased via voluntarily induced perceptual reversals (top-down control), akin to the RN (Pitts et al., 2008). Such studies could further support the notion that these auditory components are analogs of their visual predecessors, and could help integrate the auditory modality into discussions of consciousness and awareness.

As mentioned in prior sections of the discussion, many analyses could be performed given the data collected from the present experiment. Some of these analyses could include but are not limited to: Using multivariate pattern classification analysis to decode the neural correlates of perceiving “Laurel” versus perceiving “Yanny”, testing whether there are any correlations between the neural correlates found and the semitone frequency of stimuli used in experimental blocks, and comparing the MMN-like first negativity’s response amplitude resulting from a perceptual switch to “ Laurel” to the response amplitude resulting from a perceptual switch to “Yanny”. As it has been found that the response amplitude of the MMN is increased in the presence of a familiarly deviant language stimulus; a possible increase may be found resulting from perceptual switches to “Laurel” given that “Laurel” is a more common name and word in the English language (Pulvermüller & Shtyrov, 2006 ; Shtyrov et al., 2005 ; Shtyrov & Pulvermüller, 2007).

Another enticing avenue of research would be to attempt to determine more about the relationship between the RN and the VAN. How might the mechanisms of these two components be related to one another? If the RN and aRN were found to play a role in early awareness as evidence suggests the VAN and AAN do (Eklund & Wiens, 2018 ; Eklund & Wiens, 2019), such a study could eventually help to detail what the underlying mechanisms of the LPC may be.

Intaite et al. (2010) hypothesized that if the RN was a subtype of the VAN, then the magnitude of the RN would increase correlative to the amount of the presented stimulus content that were perceived to perceptually reverse. To test this, they repeatedly presented one unilateral Necker cube and compared the magnitude of the RN in these conditions to repeatedly presented instances of two side by side bilateral Necker cubes. The results refuted their hypothesis: the RN magnitude did not correlate with the content of awareness presented to participants.

There are other ways to explore possible links between the RN and VAN. The RN commonly appears to index perceptual changes of an unchanging bistable stimulus. However, it has also been proposed that the RN can be elicited by exogenously induced reversals (Kornmeier & Bach, 2006). Assuming that this is in fact the case, a future experiment could make use of this attribute of the RN by presenting slightly disambiguated visual bistable stimuli to participants at different thresholds of visibility in an intermittent response paradigm. For example, two slightly disambiguated Necker cubes (i.e.; one cube with the upper-most square slightly shaded to the foreground and the other cube with the bottom-most square slightly shaded to the foreground) could be intermittently presented to participants at a high threshold of visibility. Participants would report their perception, and perceptual reversals would be exogenously induced by the two Necker cubes based on their sequence of presentation: this would be called the “seen” condition.

The same experimental procedure and stimulus presentation sequence could then be implemented again, however this time with an “unseen” threshold of visibility; whereby the stimuli are not visible to participant's awareness. Given that the stimuli are presented in a sequence and are also disambiguated, the experimenters would know when perceptual reversals would occur and what the participants are viewing at all times, even during the "unseen" condition. Experimenters could then compare data from the "seen" versus "unseen" Necker cube conditions, and look for the RN in time-locked EEG data. The theory is that if the RN were to be found in the "unseen" condition or both the "seen" and "unseen" conditions; this evidence would support the notion that the RN is unrelated to the VAN as the VAN is only correlated with stimuli that participants are aware of. However, if the RN were to only be found in the "seen" condition with a similar scalp topography and timing as the VAN, this could be due evidence that the components are related. A study similar in form to this one could further interpret the RN's correlation with awareness, as the paradigm employed by Intaite et al. (2010) only compared visibly evident representations of stimuli.

Future research should also look into other EEG paradigms using auditory bistable language stimuli. The current study isolated an auditory processing sequence that implements all three effects: the MMN-like first negativity, the aRN, and the aLPC. Could it be possible that other auditory language stimuli bring about a phonetically induced MMN-like negativity, followed by the aRN and aLPC? This is an interesting question to ask with other perceptually ambiguous auditory language stimuli that have also gone viral. Perhaps similar EEG paradigms would similarly find that a phonetically induced MMN would be evident, subsequently followed by an aRN and aLPC. If the findings of the current study were further replicated by other experiments using auditory bistable language stimuli, the MMN could be argued to be elicited by one single stimulus and by changes in perception.

As mentioned previously, the aLPC of the present study was much longer lasting (appearing from ~400 - 1000 milliseconds) as compared to that of Davidson & Pitts (~320 - 380 milliseconds). Future studies should attempt to interpret if this longer duration is a signature of language auditory aLPC components, as language stimuli are interpreted differently from bistable streaming stimuli.

Two possible examples of perceptually ambiguous auditory language stimuli that could be implemented in future research can be listened to at the links below:

“Green Needle & Brainstorm” (The ATTIC DWELLERS, 2018)
https://www.youtube.com/watch?v=1okD66RmktA

“That is embarrassing” (MoneyBags73, 2021)
https://www.youtube.com/watch?v=7hQV77BP-30

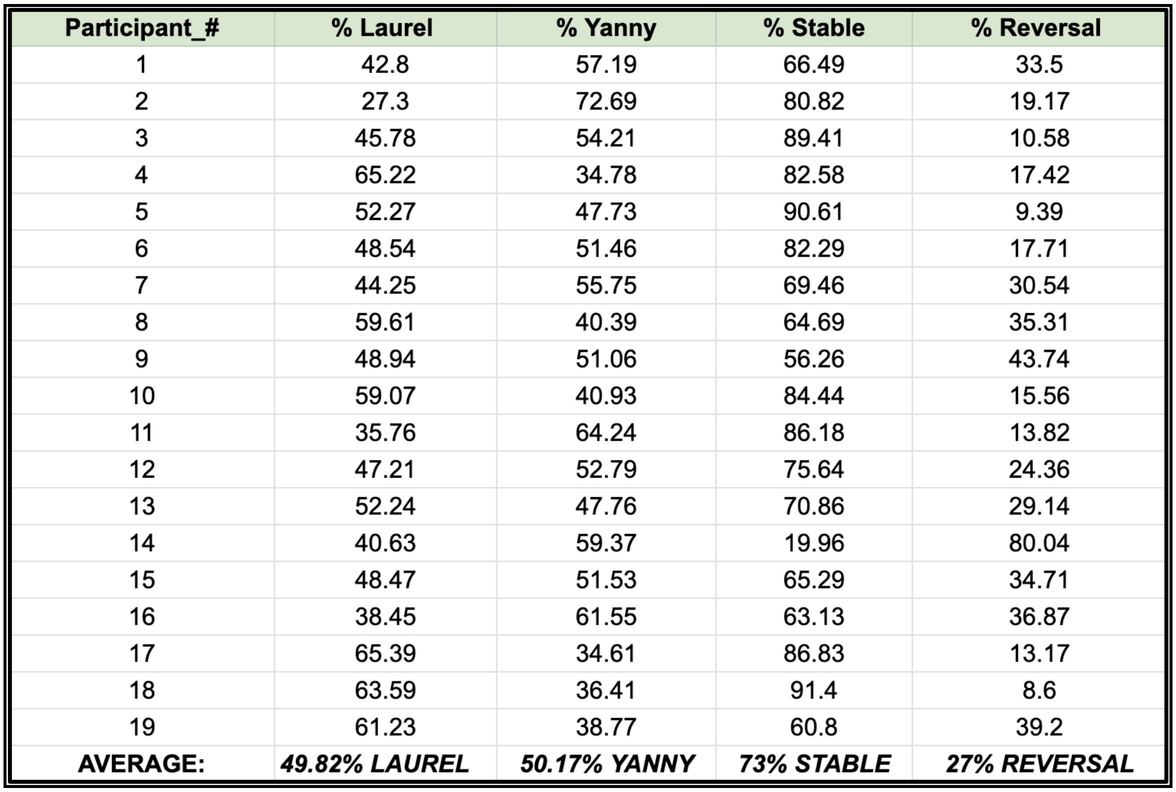
4.11 Conclusion

The results of the present study revealed three components located during perceptual reversals: a Mismatch-Negativity like component (MMN), followed by an auditory Reversal Negativity (aRN) and an auditory Late Positive Complex (aLPC). These results replicate and support recent evidence suggesting that the aRN and aLPC are auditory analogs of the widely studied visual reversal negativity (RN) and Late Positive Complex (LPC). The results further support previous theories that the MMN and aRN are distinct components (Davidson & Pitts, 2014). Results also propose that the MMN may result from endogenously perceived phonetic changes of a single auditory language stimulus; a novel instance of the MMN if so.

The present study was the first to explore the EEG neural correlates of a bistable auditory language stimulus. Its findings importantly further validated that the RN and the LPC components have auditory analogs across modalities, having been reflected in auditory language stimuli, auditory streaming stimuli, as well as in a myriad of studies examining visual bistability (Abdallah & Brooks, 2020 ; Britz et al., 2009 ; Britz & Pitts, 2011 ; Davidson & Pitts, 2014 ; Higgins et al., 2020 ; Intaite et al., 2010 ; Kornmeier & Bach, 2004, 2005; Kornmeier et al., 2007; Pitts, 2007, 2009). These results also help to detail with accurate timing and polarity events that occur in the brain related to changes in conscious perception of an unchanging stimulus.

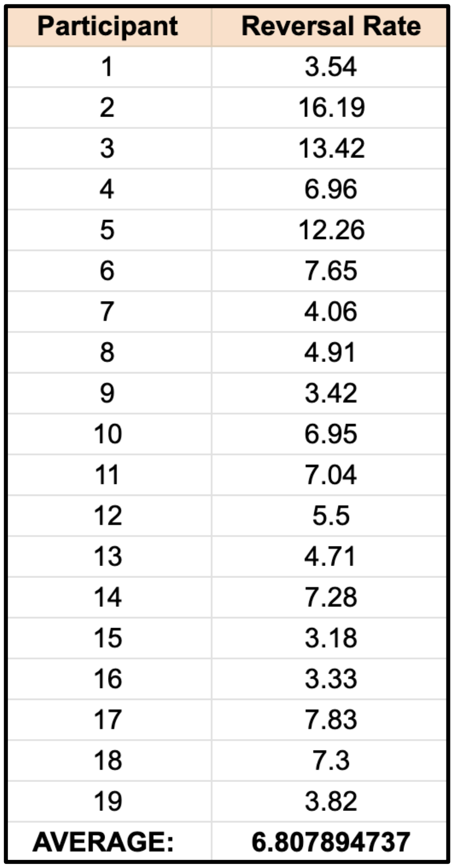
In the bigger picture, the findings of this study are a small contribution to a body of work formed by scientists who have long been dreaming to unveil the mechanisms behind consciousness, awareness, and reality.

Appendix A



Participant_#	% Laurel	% Yanny	% Stable	% Reversal
1	42.8	57.19	66.49	33.5
2	27.3	72.69	80.82	19.17
3	45.78	54.21	89.41	10.58
4	65.22	34.78	82.58	17.42
5	52.27	47.73	90.61	9.39
6	48.54	51.46	82.29	17.71
7	44.25	55.75	69.46	30.54
8	59.61	40.39	64.69	35.31
9	48.94	51.06	56.26	43.74
10	59.07	40.93	84.44	15.56
11	35.76	64.24	86.18	13.82
12	47.21	52.79	75.64	24.36
13	52.24	47.76	70.86	29.14
14	40.63	59.37	19.96	80.04
15	48.47	51.53	65.29	34.71
16	38.45	61.55	63.13	36.87
17	65.39	34.61	86.83	13.17
18	63.59	36.41	91.4	8.6
19	61.23	38.77	60.8	39.2
AVERAGE:	***49.82% LAUREL***	***50.17% YANNY***	***73% STABLE***	***27% REVERSAL***

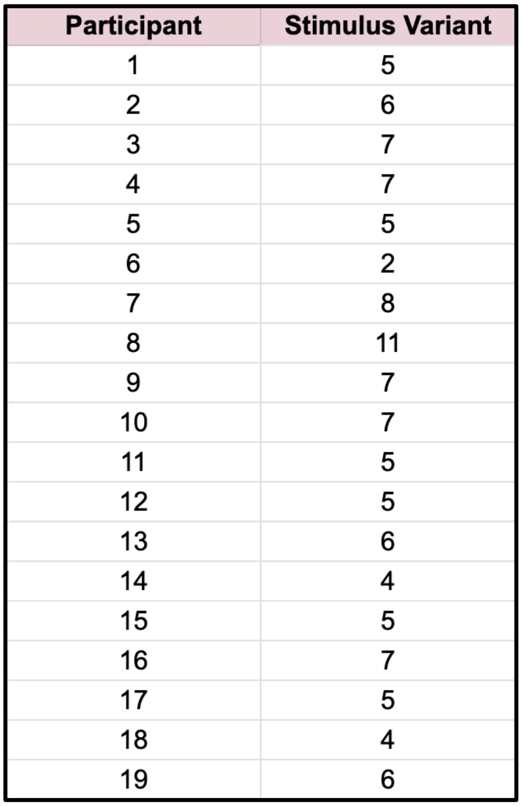
Figure A1 Participant Percent Perception Table.
Average percentages at which all four conditions were perceived for each participant analyzed for all conditions (N = 19). Mean percentages are available on the bottom row.



Participant	Reversal Rate
1	3.54
2	16.19
3	13.42
4	6.96
5	12.26
6	7.65
7	4.06
8	4.91
9	3.42
10	6.95
11	7.04
12	5.5
13	4.71
14	7.28
15	3.18
16	3.33
17	7.83
18	7.3
19	3.82
AVERAGE:	**6.807894737**

Figure A2 Reversal Rates Table.

Table reporting the reversal rates for each individual participant and the average reversal rate across participants. A reversal rate reports the average number of stable trials that took place before a reversal occurred for a participant. A larger reversal rate reflects that it took longer to perceive a perceptual reversal, whereas a smaller reversal rate reflects that it took shorter to perceive a perceptual reversal.



Participant	Stimulus Variant
1	5
2	6
3	7
4	7
5	5
6	2
7	8
8	11
9	7
10	7
11	5
12	5
13	6
14	4
15	5
16	7
17	5
18	4
19	6

Figure A3 Participant Stimulus Variant Table. The 19 participants who were included in all analyzed conditions of the EEG experiment and their respective semitone frequencies used during the study.

Bibliography

Abdallah, D., & Brooks, J. L. (2020). Response dependence of reversal-related ERP components in perception of ambiguous figures. *Psychophysiology*, *57*(12), e13685. https://doi.org/10.1111/psyp.13685

Acoustical Society of America (ASA). (2014, October 28). Ancient auditory illusions reflected in prehistoric art?. *ScienceDaily*. Retrieved May 12, 2021 from www.sciencedaily.com/releases/2014/10/141028114714.htm

Alho K. (1995). Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. *Ear and hearing*, *16*(1), 38–51. https://doi.org/10.1097/00003446-199502000-00004

Baars, B. J., Geld, N., & Kozma, R. (2021). Global workspace theory (GWT) and prefrontal cortex: Recent developments. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.749868

Baars B. J. (2005). Global workspace theory of consciousness: toward a cognitive neuroscience of human experience. *Progress in brain research*, *150*, 45–53. https://doi.org/10.1016/S0079-6123(05)50004-9

Bartha-Doering, L., Deuster, D., Giordano, V., am Zehnhoff-Dinnesen, A., & Dobel, C. (2015). A systematic review of the mismatch negativity as an index for auditory sensory memory: From basic research to clinical and developmental perspectives. *Psychophysiology*, *52*(9), 1115–1130. https://doi.org/10.1111/psyp.12459

Billig, A. J., Davis, M. H., & Carlyon, R. P. (2018). Neural Decoding of Bistable Sounds Reveals an Effect of Intention on Perceptual Organization. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *38*(11), 2844–2853. https://doi.org/10.1523/JNEUROSCI.3022-17.2018

Britz, J., Landis, T., & Michel, C. M. (2009). Right parietal brain activity precedes perceptual alternation of bistable stimuli. *Cerebral cortex (New York, N.Y. : 1991)*, *19*(1), 55–65. https://doi.org/10.1093/cercor/bhn056

Britz, J., & Pitts, M. A. (2011). Perceptual reversals during binocular rivalry: ERP components and their concomitant source differences. *Psychophysiology*, *48*(11), 1490–1499. https://doi.org/10.1111/j.1469-8986.2011.01222.x

Britz, J., Pitts, M. A., & Michel, C. M. (2011). Right parietal brain activity precedes perceptual alternation during binocular rivalry. *Human brain mapping*, *32*(9), 1432–1442. https://doi.org/10.1002/hbm.21117

Casale, J., Kandle, P. F., Murray, I., & Murr, N. (2021). Physiology, Cochlear Function. In *StatPearls*. StatPearls Publishing.

Cohen, L. (1959). Rate of apparent change of a Necker cube as a function of prior stimulation. *The American Journal of Psychology, 72,* 327–344. https://doi.org/10.2307/1420037

Corum, J., & Huang, J., & Katz, J. (2018). We Made a Tool So You Can Hear Both Yanny and Laurel. *The New York Times*, www.nytimes.com/interactive/2018/05/16/upshot/audio-clip-yanny-laurel-debate.html.

Davidson, G. D., & Pitts, M. A. (2014). Auditory event-related potentials associated with perceptual reversals of bistable pitch motion. *Frontiers in human neuroscience*, *8*, 572. https://doi.org/10.3389/fnhum.2014.00572

Dehaene, S., & Changeux, J. P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron*, *70*(2), 200–227. https://doi.org/10.1016/j.neuron.2011.03.018

Dien, J., Spencer, K. M., & Donchin, E. (2004). Parsing the late positive complex: mental chronometry and the ERP components that inhabit the neighborhood of the P300. *Psychophysiology*, *41*(5), 665–678. https://doi.org/10.1111/j.1469-8986.2004.00193.x

Duffy, F. H., Albert, M. S., McAnulty, G., & Garvey, A. J. (1984). Age-related differences in brain electrical activity of healthy subjects. *Annals of neurology*, *16*(4), 430–438. https://doi.org/10.1002/ana.410160403

Eklund, R., & Wiens, S. (2018). Visual awareness negativity is an early neural correlate of awareness: A preregistered study with two Gabor sizes.

Cognitive, affective & behavioral neuroscience, *18*(1), 176–188. https://doi.org/10.3758/s13415-018-0562-z

Eklund, R., & Wiens, S. (2019). Auditory awareness negativity is an electrophysiological correlate of awareness in an auditory threshold task. *Consciousness and cognition*, *71*, 70–78. https://doi.org/10.1016/j.concog.2019.03.008

El-Hai, J. (n.d.). *America's First Pop Psychologist.* Wonders & Marvels. https://www.wondersandmarvels.com/2012/10/americas-first-pop-psychologist.html

Fairhall, S. L., & Macaluso, E. (2009). Spatial attention can modulate audiovisual integration at multiple cortical and subcortical sites. *The European journal of neuroscience*, *29*(6), 1247–1257. https://doi.org/10.1111/j.1460-9568.2009.06688.x

Fries P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in cognitive sciences*, *9*(10), 474–480. https://doi.org/10.1016/j.tics.2005.08.011

Gajanan, M. (2018). The Surprising Origin of the 'Yanny' or 'Laurel' Clip That's Tearing the Internet Apart. *Time Magazine,* https://time.com/5279727/yanny-laurel-original-clip/

Gamble, M. L., & Luck, S. J. (2011). N2ac: an ERP component associated with the focusing of attention within an auditory scene. *Psychophysiology*, *48*(8), 1057–1068. https://doi.org/10.1111/j.1469-8986.2010.01172.x

Girgus, J. J., Rock, I., & Egatz, R. (1977). The effect of knowledge of reversibility on the reversibility of ambiguous figures. *Perception & Psychophysics*, *22*(6), 550–556. https://doi.org/10.3758/BF03198762

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: a critical tutorial review. *Psychophysiology*, *48*(12), 1711–1725. https://doi.org/10.1111/j.1469-8986.2011.01273.x

Guardian News. (2018, May 16) *Yanny or Laurel video: which name do you hear?* [Video]. Youtube. https://www.youtube.com/watch?v=7X_WvGAhMlQ

Hickey, C., McDonald, J. J., & Theeuwes, J. (2006). Electrophysiological evidence of the capture of visual attention. *Journal of cognitive neuroscience*, *18*(4), 604–613. https://doi.org/10.1162/jocn.2006.18.4.604

Higgins, N. C., Little, D. F., Yerkes, B. D., Nave, K. M., Kuruvilla-Mathew, A., Elhilali, M., & Snyder, J. S. (2020). Neural correlates of perceptual switching while listening to bistable auditory streaming stimuli. *NeuroImage*, *204*, 116220. https://doi.org/10.1016/j.neuroimage.2019.116220

Howe, C. Q., & Purves, D. (2010). *Perceiving geometry: Geometrical illusions explained by natural scene statistics*. Springer.

Intaite, M., Koivisto, M., Ruksenas, O., & Revonsuo, A. (2010). Reversal negativity and bistable stimuli: Attention, awareness, or something else?. *Brain and cognition*, *74*(1), 24–34. https://doi.org/10.1016/j.bandc.2010.06.002

Kitaoka, A. (2007). *Schröder's staircase* [Digital Image]. The latest works 25. http://www.psy.ritsumei.ac.jp/~akitaoka/saishin25e.html

Kitaoka, A. [@AkiyoshiKitaoka]. (2018, February 24). Motion assimilation (motion capture) demo using the Rotating snakes illusion (optimized Fraser-Wilcox illusion): The upper-left and lower-right red rings appear to rotate counterclockwise whereas the upper-right and the lower-left ones clockwise. [Tweet]. Twitter. https://twitter.com/akiyoshikitaoka/status/967563455250153472

Koivisto, M., & Revonsuo, A. (2010). Event-related brain potential correlates of visual awareness. *Neuroscience and biobehavioral reviews*, *34*(6), 922–934. https://doi.org/10.1016/j.neubiorev.2009.12.002

Koivisto, M., Salminen-Vaparanta, N., Grassini, S., & Revonsuo, A. (2016). Subjective visual awareness emerges prior to P3. *The European journal of neuroscience*, *43*(12), 1601–1611. https://doi.org/10.1111/ejn.13264

Kornmeier, J., & Bach, M. (2004). Early neural activity in Necker-cube reversal: evidence for low-level processing of a gestalt phenomenon. *Psychophysiology*, *41*(1), 1–8. https://doi.org/10.1046/j.1469-8986.2003.00126.x

Kornmeier, J., & Bach, M. (2005). The Necker cube--an ambiguous figure disambiguated in early visual processing. *Vision research*, *45*(8), 955–960. https://doi.org/10.1016/j.visres.2004.10.006

Kornmeier, J., & Bach, M. (2006). Bistable perception -- along the processing chain from ambiguous visual input to a stable percept. *International journal of psychophysiology : official journal of the International Organization of Psychophysiology*, *62*(2), 345–349. https://doi.org/10.1016/j.ijpsycho.2006.04.007

Kornmeier, J., & Bach, M. (2012). Ambiguous figures - what happens in the brain when perception changes but not the stimulus. *Frontiers in human neuroscience*, *6*, 51. https://doi.org/10.3389/fnhum.2012.00051

Kornmeier, J., Ehm, W., Bigalke, H., & Bach, M. (2007). Discontinuous presentation of ambiguous figures: how interstimulus-interval durations affect reversal dynamics and ERPs. *Psychophysiology*, *44*(4), 552–560. https://doi.org/10.1111/j.1469-8986.2007.00525.x

Kornmeier, J., Hein, C. M., & Bach, M. (2009). Multistable perception: when bottom-up and top-down coincide. *Brain and cognition*, *69*(1), 138–147. https://doi.org/10.1016/j.bandc.2008.06.005

Lamme, V. A., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in neurosciences*, *23*(11), 571–579. https://doi.org/10.1016/s0166-2236(00)01657-x

Lamy, D., Salti, M., & Bar-Haim, Y. (2009). Neural correlates of subjective awareness and unconscious processing: an ERP study. *Journal of cognitive neuroscience*, *21*(7), 1435–1446. https://doi.org/10.1162/jocn.2009.21064

Long, G. M., & Toppino, T. C. (2004). Enduring interest in perceptual ambiguity:

alternating views of reversible figures. *Psychological bulletin*, *130*(5), 748–768. https://doi.org/10.1037/0033-2909.130.5.748

Luo, L. (2015). *Principles of Neurobiology*. Garland Science.

Maksim & Boivie. (2008). *Necker cube and impossible cube* [Digital Image]. Wikimedia Commons.

https://en.wikipedia.org/wiki/File:Necker_cube_and_impossible_cube.svg

Mashour, G. A., Roelfsema, P., Changeux, J. P., & Dehaene, S. (2020). Conscious Processing and the Global Neuronal Workspace Hypothesis. *Neuron*, *105*(5), 776–798. https://doi.org/10.1016/j.neuron.2020.01.026

MoneyBags73. (2021, June 29) *Incredible Example Of The Hear What You Want Phenomenon!!!* [Video]. Youtube. https://www.youtube.com/watch?v=7hQV77BP-30

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R. J., Luuk, A., Allik, J., Sinkkonen, J., & Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, *385*(6615), 432–434. https://doi.org/10.1038/385432a0

Näätänen, R., Paavilainen, P., & Reinikainen, K. (1989). Do event-related potentials to infrequent decrements in duration of auditory stimuli demonstrate a memory trace in man?. *Neuroscience letters*, *107*(1-3), 347–352. https://doi.org/10.1016/0304-3940(89)90844-6

Nunez P. L. (1981). A study of origins of the time dependencies of scalp EEG: i--theoretical basis. *IEEE transactions on biomedical engineering*, *28*(3), 271–280. https://doi.org/10.1109/tbme.1981.324700

Ojanen, V., Revonsuo, A., & Sams, M. (2003). Visual awareness of low-contrast stimuli is reflected in event-related brain potentials. *Psychophysiology*, *40*(2), 192–197. https://doi.org/10.1111/1469-8986.00021

Paavilainen, P., Alho, K., Reinikainen, K., Sams, M., & Näätänen, R. (1991). Right hemisphere dominance of different mismatch negativities.

Electroencephalography and clinical neurophysiology, *78*(6), 466–479. https://doi.org/10.1016/0013-4694(91)90064-b

Pitts, M. A., Gavin, W. J., & Nerger, J. L. (2008). Early top-down influences on bistable perception revealed by event-related potentials. *Brain and cognition*, *67*(1), 11–24. https://doi.org/10.1016/j.bandc.2007.10.004

Pitts, M. A., Martínez, A., Stalmaster, C., Nerger, J. L., & Hillyard, S. A. (2009). Neural generators of ERPs linked with Necker cube reversals.

Psychophysiology, *46*(4), 694–702. https://doi.org/10.1111/j.1469-8986.2009.00822.x

Pitts, M. A., Nerger, J. L., & Davis, T. J. (2007). Electrophysiological correlates of perceptual reversals for three different types of multistable images. *Journal of vision*, *7*(1), 6. https://doi.org/10.1167/7.1.6

Pulvermüller, F., & Shtyrov, Y. (2006). Language outside the focus of attention: the mismatch negativity as a tool for studying higher cognitive processes. *Progress in neurobiology*, *79*(1), 49–71. https://doi.org/10.1016/j.pneurobio.2006.04.004

Purves, D., Morgenstern, Y., & Wojtach, W. T. (2015). Perception and Reality: Why a Wholly Empirical Paradigm is Needed to Understand Vision. *Frontiers in systems neuroscience*, *9*, 156. https://doi.org/10.3389/fnsys.2015.00156

Ramachandran, V., & Hirstein, W. (1997). Three Laws of Qualia. *Journal of Consciousness Studies*. 4. https://www.researchgate.net/publication/233684568_Three_laws_of_qualia_What_neurology_tells_us_about_the_biological_functions_of_consciousness

Robitaille, N., & Jolicoeur, P. (2006). Fundamental properties of the N2pc as an index of spatial attention: effects of masking. *Canadian journal of experimental psychology = Revue canadienne de psychologie experimentale*, *60*(2), 101–111. https://doi.org/10.1037/cjep2006011

Salti, M., Bar-Haim, Y., & Lamy, D. (2012). The P3 component of the ERP reflects conscious perception, not confidence. *Consciousness and cognition*, *21*(2), 961–968. https://doi.org/10.1016/j.concog.2012.01.012

Sandberg, J., Hansson, M., & Lindgren, M. (2005). Detecting MMN in Infants EEG with Singular Value Decomposition. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, *2005*, 4227–4230. https://doi.org/10.1109/IEMBS.2005.1615397

Shtyrov, Y., Kujala, T., Palva, S., Ilmoniemi, R. J., & Näätänen, R. (2000). Discrimination of speech and of complex nonspeech sounds of different temporal structure in the left and right cerebral hemispheres. *NeuroImage*, *12*(6), 657–663. https://doi.org/10.1006/nimg.2000.0646

Shtyrov, Y., Pihko, E., & Pulvermüller, F. (2005). Determinants of dominance: is language laterality explained by physical or linguistic features of speech?. *NeuroImage*, *27*(1), 37–47. https://doi.org/10.1016/j.neuroimage.2005.02.003

Shtyrov, Y., & Pulvermüller, F. (2007). Language in the mismatch negativity design. *Journal of Psychophysiology*, *21*(3-4), 176–187. https://doi.org/10.1027/0269-8803.21.34.176

Slotnick, S. D., & Yantis, S. (2005). Common neural substrates for the control and effects of visual attention and perceptual bistability. *Brain research. Cognitive brain research*, *24*(1), 97–108. https://doi.org/10.1016/j.cogbrainres.2004.12.008

Stevo-88. (2006). Necker's cube. [Image]. Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Necker%27s_cube.svg

Stokes, T. A., Welk, A. K., Zielinska, O. A., & Gillan, D. J. (2017). The oddball effect and inattentional blindness: How unexpected events influence our perceptions of Time. *Proceedings of the Human Factors and Ergonomics*

Society Annual Meeting, *61*(1), 1753–1757. https://doi.org/10.1177/1541931213601919

Tannenbaum A. S. (2001). The sense of consciousness. *Journal of theoretical biology*, *211*(4), 377–391. https://doi.org/10.1006/jtbi.2001.2355

The ATTIC DWELLERS. (2018, May 17) *BRAINSTORM -OR- GREEN NEEDLE (U HEAR WHAT YOU THINK!)* [Video]. Youtube. https://www.youtube.com/watch?v=1okD66RmktA

Thompson, C. (2016). How The Phonograph Changed Music Forever. *Smithsonian Magazine*. https://www.smithsonianmag.com/arts-culture/phonograph-changed-music-forever-180957677

Toppino, T. C., & Long, G. M. (1987). Selective adaptation with reversible figures: don't change that channel. *Perception & psychophysics*, *42*(1), 37–48. https://doi.org/10.3758/bf03211512

Tsuchiya, N., & van Boxtel, J. (2013). Introduction to research topic: attention and consciousness in different senses. *Frontiers in psychology*, *4*, 249. https://doi.org/10.3389/fpsyg.2013.00249

Urakawa, T., Nagano, K., Matsumoto, Y., & Araki, O. (2018). Enhancement of a genuine visual mismatch negativity correlates with the facilitation of perceptual alternation of a bistable image. *Neuroreport*, *29*(13), 1104–1108. https://doi.org/10.1097/WNR.0000000000001079

van Ee, R., van Dam, L. C., & Brouwer, G. J. (2005). Voluntary control and the dynamics of perceptual bi-stability. *Vision research*, *45*(1), 41–55. https://doi.org/10.1016/j.visres.2004.07.030

Vercillo, K. (2012). *Rubin Vase* [Digital Image]. HubPages. https://discover.hubpages.com/art/Two-Faces-or-a-Vase-10-Simple-but-Wonderful-Optical-Illusions

Vetter, D. E. (2008, January 31). *How do the hammer, anvil and stirrup bones amplify sound into the inner ear?* Scientific American. Retrieved November 16, 2021, from https://www.scientificamerican.com/article/experts-how-do-the-hammer-anvil-a/.

Ward, J. (2020). The Student's Guide to Cognitive Neuroscience. Fourth edition. Routledge, Taylor & Francis Group.

Warrier, C., Wong, P., Penhune, V., Zatorre, R., Parrish, T., Abrams, D., & Kraus, N. (2009). Relating structure to function: Heschl's gyrus and acoustic processing. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *29*(1), 61–69. https://doi.org/10.1523/JNEUROSCI.3489-08.2009

Wilenius, M. E., & Revonsuo, A. T. (2007). Timing of the earliest ERP correlate of visual awareness. *Psychophysiology*, *44*(5), 703–710. https://doi.org/10.1111/j.1469-8986.2007.00546.x

Xi, J., Zhang, L., Shu, H., Zhang, Y., & Li, P. (2010). Categorical perception of lexical tones in Chinese revealed by mismatch negativity. *Neuroscience*, *170*(1), 223–231. https://doi.org/10.1016/j.neuroscience.2010.06.077

Xu, X., Hanganu-Opatz, I. L., & Bieler, M. (2020). Cross-Talk of Low-Level Sensory and High-Level Cognitive Processing: Development, Mechanisms, and Relevance for Cross-Modal Abilities of the Brain. *Frontiers in neurorobotics*, *14*, 7. https://doi.org/10.3389/fnbot.2020.00007

Yang, H., Laforge, G., Stojanoski, B., Nichols, E. S., McRae, K., & Köhler, S. (2019). Late positive complex in event-related potentials tracks memory signals when they are decision relevant. *Scientific reports*, *9*(1), 9469. https://doi.org/10.1038/s41598-019-45880-y

Yi, H. G., Leonard, M. K., & Chang, E. F. (2019). The Encoding of Speech Sounds in the Superior Temporal Gyrus. *Neuron*, *102*(6), 1096–1110. https://doi.org/10.1016/j.neuron.2019.04.023

Zevin, J. D., Datta, H., Maurer, U., Rosania, K. A., & McCandliss, B. D. (2010). Native language experience influences the topography of the mismatch negativity to speech. *Frontiers in human neuroscience*, *4*, 212. https://doi.org/10.3389/fnhum.2010.00212