A COMPARISON OF MUSIC AND PROSODIC PROCESSING IN AUTISM SPECTRUM DISORDER

AN ABSTRACT
SUBMITTED ON THE SEVENTH DAY OF APRIL 2015
TO THE PROGRAM IN LINGUISTICS
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
OF THE SCHOOL OF LIBERAL ARTS
OF TULANE UNIVERSITY
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
BY

JOHN DePRIEST

APPROVED: Harry Howard
HARRY HOWARD, Ph.D.
DIRECTOR

Judith M Maxwell
JUDITH MAXWELL, Ph.D.

ED GOLOB, Ph.D.

Abstract

Autism Spectrum Disorders (ASDs) are frequently associated with communicative impairment, regardless of IQ or mental age. The most significant feature of this impairment tends to be in the dimension of both expressive and receptive prosody, possibly due to reduced neural connectivity between disparate brain areas responsible for language. Despite extensive overlap between the auditory and structural features linking prosody and music as well as extensive shared neural resources, music listening and performance are not impaired. In fact, there is some evidence that these abilities may even be heightened in some ASD individuals. Using behavioral and EEG/ERP methods, the present study sought to investigate this dissociation. A similar electrophysiological response has been observed for both prosody and music, the Closure Positive Shift (CPS), and Music CPS, respectively. This study used language and music stimuli in order to investigate the differences between language and music processing for individuals with ASDs and neuro-typicals. While a CPS was observed for language for the ASD group, it was substantially reduced in its distribution and amplitude. Further, the presence of an offset N1 response to the onset of pauses interfered with the clarity of the CPS response. In music, no music CPS was observed, however, a sustained centrally maximal positivity was observed for both the neuro-typical and ASD groups during the phrase boundary. Additionally, the ASD group showed a similar positivity in response to phrase boundaries in the condition in which the phrase-final note was prolonged. This positivity was similar to the language CPS in duration and amplitude, and suggests similar processing responses to phrase boundaries in language and music. The positivity in response to the second condition suggests that some individuals with ASDs may indeed have heightened processing ability for music. These results support the theories of functional under-connectivity in language and local bias toward sensory features of auditory information at the expense of global prosodic processing. Possible explanations, including the presence of repetition found in music, yet generally absent in language, are considered.

A COMPARISON OF MUSIC AND PROSODIC PROCESSING IN AUTISM SPECTRUM DISORDER

A DISSERTATION
SUBMITTED ON THE SEVENTH DAY OF APRIL 2015
TO THE PROGRAM IN LINGUISTICS
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
OF THE SCHOOL OF LIBERAL ARTS
OF TULANE UNIVERSITY
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
BY

JOHN DePRIEST

APPROVED: Harry Howard
HARRY HOWARD, Ph.D.
DIRECTOR

Judith M. Maxwell
JUDITH MAXWELL, Ph.D.

ED GOLOB, Ph.D.

Acknowledgments

First, I would like to thank the subjects who participated in this study, both those with autism spectrum disorders, and those without. Without participants, no research on humans could be conducted at all, and without the subjects who provided their time and attention to participation in this study, this project would not have been possible.

Foremost, I would like to thank my wife, Sara, for her astounding and never-ending love and support. She graciously tolerated many late hours and stressful days, even being willing to uproot and move to Berlin while I designed and conducted this research. She is an inspiration for hard work and unconditional love, and she brings joy to my life everyday.

None of this project would have been possible if not for my supervisors and committee who steered the research from its outset, gave hours of time reading previous versions, and provided the resources with which to conduct this study. My advisor at Tulane, Harry Howard, has also played a hugely important role in this project. He inspired me to use EEG methods to answer linguistic questions from the outset, and gave me the freedom to pursue my interests in the manner of my choosing. For that, I am endlessly grateful. My supervisor in Berlin, Stefan Koelsch, welcomed me into his lab and encouraged me to expand my scientific horizons to ask meaningful and well-directed questions that expand the boundaries of human knowledge, and it must be said that any failing in this regard in the present work is mine alone. Without Stefan's contributions, this project would have been vastly different, and of vastly lower quality. Judie Maxwell served as my de facto advisor during Harry's sabbatical, and has been a tremendous inspiration. Not only is she a linguist of the highest order, but she also possesses a determination to apply linguistics to serve humanity, particularly those in need, whether that need be help finding a voice, or finding a way to make that voice heard. Ed Golob made numerous suggestions to earlier drafts of this project which

have improved its quality tremendously. His encouragement was invaluable. His dedication to science, and to methods of carefully and precisely pushing the edges of what we know has made me incorporate these values as my own.

Many other professors and researchers have spent time helping me in my quest to conduct original and meaningful research. Nike Orie, who first inspired my interest in phonetics, phonology, and prosody, has been a constant help from the beginning, both as a professor, and head of the Linguistics Program, as well as as a friend. Karsten Steinhauer has also been a tremendous help. Not only did he help discover the CPS, without which this study would not have been possible, but he also believed in this project enough to contribute thoughtful comments and suggestions during its design, as well as contributing his time, expertise, and assistance during data analysis in Montreal. For this I am grateful. Sebastian Jentschke also made many comments on earlier versions of this project, and his methodological expertise, particularly his wizardry with scripting, were a great help in conducting early analysis of this data. I am grateful for his training, and to call him a friend. Many other professors helped with details of the project and inspiration in other ways, including Isabel Dziobek, who made valuable comments on details of working with autism spectrum participants, and Nick Spitzer, who has years of experience working with language and music, and who gave me a window into the cultural and ethnographic side of this world.

Of course, the organizations that supported me through coursework and research have made this entire project possible. This dissertation could not have been completed without the generous financial support of Tulane University, the Freie Universität Berlin, and the Choctaw Nation of Oklahoma.

This project was also aided tremendously by the contributions of the unsung heroes of academia: graduate students, both at Tulane, the Freie Universität Berlin, and elsewhere. These include Shayra Garcia, who has been a contemporary and primary commiserator throughout my graduate tenure; Stavros Skouras, who helped teach me MatLab, and made helpful comments during the development of this project; Moritz Lehne, for always taking the time to get to the heart of an issue, including early issues with the design of this experiment; Laura Hahn, for help developing stimuli, and numerous discussions on methods and linguistic research; Anne Märtin, who provided the voice for the language stimuli; and Ariadne Saenz, who, as a composer, gave helpful review of the music stimuli. The other members of the Neuroscience of Music lab at the FU were also helpful in many ways, including Liila Taruffi, Philipp Engel, Susanna Greifeneder, and Aleksandra Gulka. Anastasia Glushko deserves special mention. Her contributions to the recruitment and testing of subjects, and to data analysis were tremendously helpful. She has an exciting career to look forward to.

Many friends provided invaluable sounding boards, comments, and various other types of support without which this study would have suffered. Primary among these is Katharina Jenderny, who graciously opened her home to my wife and me in Berlin after having known us for three days. She became like a family member and helped with translation, idea refinement, and caffeination. Others include Philipp Peters, Jula Pötter, Martin Mittermaier, Sophie Caporal, Constanze Flamme, Frank Flöthmann, Marcel Türkowsky, Elise Florenty, Hiwa Michaeli, Pavel Jansta, and Nabil Atassi. There are far too many to list here, but thank you all.

Many fellow musicians in New Orleans provided inspiration to study music in the first place, and much needed relief during moments of stress. These include first and foremost Will Jordan, as well as Josh Crist, John Norwood, Michael Millet, Ron Hotstream, Blake Mogabgab, Mike Harvey, James & Lindsay Hausmann, Lauren Hemard, Asher Griffith, and dear old Ian Wood.

None of this would have been possible without the help of my family, both DePriest and Luton. My parents, Don & Sandra, have provided endless support, as have my siblings Warner, Robert, Sally, and her family, and all of the Lutons.

Contents

List of Tables

List of Figures

1 Introduction

1.1 Background

1.1.1 Autism

The history of official diagnosis of autism dates from 1943 when Leo Kanner published his seminal work entitled, *Autistic Disturbances of Affective Contact.* (Kanner, 1943) This article served to establish the initial diagnostic criteria and the basic description of the symptoms of autism. The next year, Hans Asperger published another work on the subject entitled *Die "Autistischen Psychopathen" im Kindesalter.* (Asperger, 1944) These two works formed the basis of what would later be referred to as Autism Spectrum Disorders (ASDs), since the two papers described slightly different manifestations of a similar range of symptoms. Today, the disorder described by Kanner and that described by Asperger are distinctly classified, particularly in terms of degree of language delay and impairment, but they overlap, and are generally considered to be part of the same spectrum.

Since that time, interest in the condition has swelled considerably, partly due to the fact that the disorder is far more common than originally supposed. What was originally thought to appear in 5:10,000 children is now diagnosed at a rate closer to 1:110 (Amaral, et. al., 2011; Lai, et. al., 2012), yet there is substantial variability in diagnostic rates in general, as well as across national boundaries due to levels of awareness and medical service availability. (for a review, see Elsabbagh, et al., 2012) It is almost certain that the increase in diagnosed cases is due simply to increased awareness and diagnostic frequency, as opposed to the prevalence of the condition actually increasing. (Elsabbagh, 2012) In either case, the disorder has become common enough to warrant global interest and millions of dollars of allocated resources for research, diagnosis, and treatment, including the Combating Autism Act signed by President Bush in 2006. An article for new researchers in autism remarks that "Kanner's paper was only referenced 34 times between 1943 and 1954. By contrast, it was referenced nearly 140 times in 2009 alone." (p. 4, Amaral, et al., 2011)

Neural Organization in Autism

Previously, it was thought that autism was caused by purely social and developmental factors, including a theory that cold, distant, 'refrigerator' mothers were the cause of autism. This idea was dispelled by the recognition that the co-morbidity of neural symptoms (such as epilepsy) signals an underlying neurological disorder unlikely to be caused by inattention. (Rimland, 1964; Sacks, 1995) Numerous studies investigating twin and family prevalence of autism have been conducted in recent years, yet it remains unclear whether autism is fully heritable, and it seems likely that environmental factors do indeed play a role in the etiology of autism. (Hallmayer, et al., 2011; Sandin, et al., 2014; Schendel, et al., 2014) Regardless of the cause, the discovery that autism is often associated with the presence of severe neurological impairments spurred researchers to try to identify the underlying neural structures involved in the disorder. Still, a precise neurophysiological description of the disorder remains elusive. (for a review, see Jeste, et al., 2011)

Researchers have additionally sought to establish consistent theories of cognition as manifested by this disordered neural architecture. Many of these interpretations focus on what is known as a disordered 'Theory of Mind'. This idea was largely established in the late 80s and early 90s by a team of researchers at Cambridge, including Uta Frith and Simon Baron-Cohen. (Baron-Cohen, et. al., 1985; Frith, 1989) The premise is that many individuals with ASDs are impaired in their ability to recognize that other people have separate minds and to understand that others may not share the same knowledge, desires, or interests that they themselves possess. Baron-Cohen has termed this lack of recognition of others' minds as 'mindblindness'. (Baron-Cohen, 1995) While the Theory of Mind hypothesis has declined somewhat in popularity due to the limits of its explanatory power, (Frith, & Happé, 1994; Chevallier, et. al., 2011) it remains an important description of the social abilities and limitations of individuals with ASDs, largely due to the theory's emphasis on joint attention. (Sigman, 1998; Mundy & Gomes, 1998; Ames & Fletcher-Wilson, 2010; Maljaars, et. al. 2011)

From a neurological perspective, many studies have been conducted in the last 20 years with the goal of determining what features are common to the spectrum of autistic brains. Researchers have focused on a wide variety of brain regions in an attempt to isolate the specific areas or systems that are disordered in autism, with wide ranging results. Researchers have focused on areas related to social functioning, such as the Superior Temporal Sulcus (STS), the orbitofrontal and medial prefrontal cortices, and the amygdala. (Baron-Cohen, 1995; Groen, et al., 2008) Research related to language deficits in autism has identified an over-reliance on Wernicke's area, (posterior third of the Superior Temporal Gyrus, STG), (Just, et al., 2004; Groen, et al., 2008), decreased and even reversed laterality of Broca's area (Inferior Frontal Gyrus, IFG) (Herbert, et al., 2002; Hodge, et al., 2010) and the cerebellum. (Hodge, et al., 2010) One study (Herbert, et al., 2002) found that a participant population with ASDs had right hemisphere inferior frontal gyri that were 27% larger than that found in the left hemisphere, contrary to right-handed neuro-typical individuals who consistently show hemispheric asymmetry with larger left hemisphere language areas. Other studies have provided further confirmation of the IFG asymmetry in ASD individuals. De Fosse, et al. (2004) observed the same reversed asymmetry, however, only for those ASD participants with linguistic impairment.

In addition to the idea that single areas of the brain are disordered, other research has taken a more systems level or network approach to describing neural abnormality in ASDs. The mirror-neuron (MN) system, for example, has received substantial attention regarding its deficit in autism. (i.e. Williams, et. al., 2001; Wan et al., 2010; see Perkins, et. al., 2010 for a review):

> “Although there is no reliable neurophysiological marker associated with ASDs, dysfunction of the parieto-frontal mirror neuron system has been suggested as a disturbance linked to the disorder. Mirror neurons (MNs) are visuomotor neurons which discharge both when performing and observing a goal directed action. Research suggests MNs may have a role in imitation, empathy, theory of mind and language.”
>
> (p. 1239, Perkins, et. al., 2010)

The network of mirror-neurons is believed to be distributed throughout many of the areas mentioned in the preceding paragraphs (i.e., STS, IFG, as well as somato-sensory cortices), and are thought to represent a link between action and language, playing a major role in embodied cognition, and theory of mind. (Oberman & Ramachandran, 2007; Le Bel, et al., 2009) While the explanatory power of this hypothesis is much broader than that of simply a ‘Theory of Mind’ deficit, it presents challenges for falsification. The functioning of the MN system in humans, and its relevance to language (and cognition as a whole) is not entirely understood. Indeed, much of what we know about the MN system is derived from research on primates, and speculative hypotheses about its role in the human brain. (Hickok, 2009)

Under-Connectivity Theory

One of the primary hypotheses regarding the neurological phenotype of ASDs is called the ‘under-connectivity theory’. Many studies over the past decade have provided evidence in support of functional and structural under-connectivity between regions of autistic brains, particularly relating to frontal-posterior connections. (Just, et. al., 2004; Just, et. al., 2012; Lai, et. al., 2012) One of the earliest papers proposing this hypothesis describes it as referring "to the under-functioning of integrative circuitry and emergent cognitive, perceptual, and motor abilities in autism." (p. 1817, Just, et. al., 2004) The authors go on to say that this under-functioning circuitry "results in a deficit of integration of information at the neural and cognitive levels." (p. 1817, Just, et. al., 2004) Taken together, this theory proposes that a long-range structural deficit in neural architecture results in a functional bandwidth limitation of information consolidation across disparate regions. This bandwidth limitation could cause cognitive impairments in nearly any dimension where a task requires integration of information from distant neural areas, particularly when the cognitive demand is high.

This hypothesis stems from studies on language, a behavior that requires high cognitive demand involving disparate regions of the brain. The results of these linguistic fMRI studies show that during visual sentence processing, autistic individuals have higher levels of activation in the posterior Superior Temporal Gyrus (STG), or Wernicke's area, and lower levels of activation in the Left Inferior Frontal Gyrus (LIFG), corresponding to both Broca's area (BA 44/45) and other nearby, but cytoarchitectonically dissimilar regions such as BA 47. (Just, et. al., 2004) The same study showed that not only did the STG and LIFG appear to be underconnected in autistic subjects, but also that the Angular Gyrus of the parietal lobe showed less activation in this group than in typically developing (TD) controls. In regards to language, the authors write that they "hypothesized a lower level of functional connectivity among the autistic participants, because a difficulty in the integrative aspects of understanding a complex sentence could well stem from a lower level of coordination and synchronization among cortical areas." (p. 1813, ibid., 2004) The authors go on to mention that this hypothesis could even serve to explain the Theory of Mind deficits in autism.

Equally important to this theory is the hypothesized corresponding over-connectivity in short-range connections found in individuals with ASDs. (Groen, et al., 2008; Just, et al., 2012) This may be due to compromised dendritic and synaptic pruning in autistic individuals during crucial developmental periods (Frith, 2003), leading to the elimination of energy demanding long-range connections as development progresses. (Lewis & Ellman, 2008) The result of this over-connectivity may explain why individuals with autism have difficulties filtering environmental input to extract signal from noise, leading to problems with selective attention, overstimulation in noisy or chaotic environments, and discomfort in situations of physical contact. (e.g. Grandin, 1996) It may also, however, explain many autistic individuals' heightened attention to, and memory of, detail in visual, auditory, and/or tactile domains. (Liss, et al., 2006; Liss, et al., 2008)

Despite this hypothesis's explanatory power and preliminary evidence providing support for under-connectivity, neurological imaging methods are not yet capable of definitively falsifying it. Long-range deficits in structural connectivity coupled with a pathological short-range over-abundance fit with many behavioral descriptions of the disorder, yet the exact definitions of long and short range are elusive, meaning that an over-, or under-abundance of medium range connections may present problems of classification, or even a fluidity of definition that is too imprecise to be useful. Also, whether a lack of un-coordinated activation in frontal/posterior regions is due to a lack of connectivity or any number of other factors, such as connectivity to different regions than those found in neurotypicals, or over-inhibition of a particular assembly of cells, remains unclear. Still, this hypothesis has the most explanatory power of the wide-ranging cognitive and structural theories of ASDs, and a growing body of evidence to support it.

Part of the difficulty with developing consistent theories to better understand and treat ASDs comes from the enormous heterogeneity of symptoms and functioning level classified under this spectrum. On the lower functioning end, sufferers can have comorbidity with severe mental retardation and epilepsy, while on the higher functioning end, individuals with this disorder have obtained Ph.D.s and pursued highly successful careers. The most famous example of an individual in this latter category is the animal scientist, author, and professor Temple Grandin, who was profiled in Oliver Sacks's book *An Anthropologist on Mars* (2005), as well as having written numerous books of her own. The title of Sacks's book refers to a quote about how foreign Grandin feels when interacting with humans, despite her range of achievements. (Sacks, 1995) While there are symptoms common to nearly all sufferers of autism, this wide disparity of functioning leads to difficulties in deciphering the causes of and specific impairments associated with ASDs.

Despite this symptomatic and diagnostic heterogeneity in the autism spectrum, one of the most commonly recognized disabilities is that of atypical language ability. In general, people with autism not only have difficulties with language processing, but also with language production. In the DSM-IV-TR, one of the diagnostic criteria for autism is:

> (2) qualitative impairments in communication as manifested by at least one of the following:
> (a) delay in, or total lack of, the development of spoken language (not accompanied by an attempt to compensate through alternative modes of communication such as gesture or mime)
> (b) in individuals with adequate speech, marked impairment in the ability to initiate or sustain a conversation with others
> (c) stereotyped and repetitive use of language or idiosyncratic language
>
> (APA, 2000)

The reasons for atypical language abilities are still not well understood, but this has not kept some of the features of language that present the greatest difficulty from being described. The most commonly addressed issue in autistic language ability, however, is that of prosody.

1.1.2 Prosody

Prosody can refer to many features of language, but is largely considered to refer to the supra-segmental acoustic elements of the speech stream. Supra-segmental features are the "aspects of speech that involve more than single consonants or vowels" including stress, tone, syllable length, rhythm, and intonation. (p. 237, Ladefoged, 2006) de Angulo, in his influential 1929 article in *Language*, defined prosody as follows: “I would like to suggest the use of the term ‘prosody’ to include all those changes which are often lumped together as accentual differences. They involve three independent factors: pitch (or tone), duration (or length, quantity), amplitude (or volume, stress, loudness, etc.).” (de Angulo, 1929) While these two definitions together are quite accurate for the general considerations of prosody, it is important to incorporate emotional or affective prosody as well.

The functions of prosody are manifold. Roach (2000) suggests that the purposes of prosody are syntactic, pragmatic, and affective, although there is overlap between these three categories. (Peppé et. al., 2007) I would add that the purpose of prosody is to draw attention, or to cue the listener to these elements of the speech stream. A speaker can specify phrase or sentence Focus, or can signal the end of a turn. A speaker can mark degree through iconic prosodic representations such as pitch or amplitude shifts, while conveying anger or joy. Regardless of the immediate use, the main purpose of prosody is to draw the attention of the listener to some meaningful or grammatical element of the speech stream.

The importance of prosody is wide-reaching, making it one of the most defining features of spoken language. Much of prosody centers around the functions of the fundamental frequency of a phrasal utterance. The f(0) refers to the lowest frequency of speech, and is what is largely responsible for the more global 'melody' of a sentence. It is also where more local functions such as vowel length, tone, and stress, which combine to create the 'melody' are manifested. (Ladefoged, 2006) These can be mapped using spectrograms and the analysis of frequency contours of a sentence is becoming common. (Beckman & Venditti, 2010)

The methodology of studying prosody is not, however, limited to sheer physical analysis. In one of the most important EEG/ERP studies on prosody in the last few decades, Steinhauer, Alter, and Friederici (1999) found that a rapid positive going electrophysiological response is associated with phrase boundaries of spoken language. In attempting to isolate what the exact features of spoken language sentence boundaries are, the authors describe the most salient and consistent acoustical features of phrase boundaries as pre-final syllable lengthening and a slight rise in pitch, followed by a pause. (ibid., 1999) They also note that the observed effect (named the Closure Positive Shift, or CPS) could be elicited without the presence of a pause at the boundary of the intonational phrase (IPh).

(following nomenclature convention from Selkirk, 1984) Additionally, the presence of a spliced condition, creating a sentence with incongruous prosodically driven syntactic interpretation led to Garden Path responses of a bi-phasic N400 and P600, demonstrating that prosody contributes to the successful parsing of sentence structure. (Steinhauer, et al., 1999) Later, similar electrophysiological effects were observed at phrase boundaries in written language signaled by commas. (Steinhauer & Friederici, 2001) This study was important because it showed that the effect of presentation modality in language studies (visual vs. auditory) did not influence the cognitive chunking of phrases, as prosody was processed covertly in the visual domain. Together, these results suggest that phrase boundary segmentation is aided by prosodic features, including a variety of acoustic cues in the auditory domain, and commas in the visual modality.

Following the initial studies on electrophysiological correlates of prosodic specification, there have been numerous studies examining the interaction of prosody with other features of language, as well as the development of prosodic processing in children. (Pannekamp, et. al. 2005; Männel & Friederici, 2009; Männel & Friederici, 2011) These include prosody and syntax (Kerkhofs, et. al., 2007; Sammler, et. al., 2010; Eckstein, et. al., 2006; Schmidt-Kassow et. al., 2009; Strelnikov, et. al., 2006; Roll, et. al., 2011) prosody and semantics (Astésano, et. al., 2004; Kotz, et. al. 2007; Paulmann & Kotz, 2008), prosody and music (Gordon, et al., 2011) and prosody and discourse, largely in the form of turn-taking (Magne, et. al., 2005; Toepel, et. al., 2007; Cowles, et. al., 2007). Emotional prosody has also been an important field of study, largely because of the frequency of psychological and neurological disorders (such as SLI and autism) in which individuals experience difficulty perceiving emotional contrast and the state of the speaker indexed thereby. (Schirmer, et. al., 2006; Kotz, et. al., 2006; Paulmann, et. al., 2008; Paulmann, et. al., 2012) There also appears to be some influence of gender, which has been found to relate to differences in emotional prosody processing ability. (Schirmer, et. al., 2002)

Other methodological strategies for measuring prosodic processing include recent pupillometric studies. Zellin, et. al. (2011) have shown that incongruent focus specification elicited by prosodic contrasts can cause an increase in pupil dilation. The hypothesized reason for this effect is based on earlier findings that increased pupil diameter (when light level is controlled for) relates to increased cognitive load processes. Presumably, incongruous prosody results in an increased demand on cognitive load, thereby causing the pupil dilation. Nearly all of the studies mentioned in this section focus on receptive prosody, which refers to one's ability to interpret prosodically specified linguistic features in a meaningful way. The reason for this focus is methodological. Controlling for movement and exact timing is necessary for neurophysiological investigation, yet expressive prosody is an equally important element in the communicative equation. Expressive prosody tends to be more appropriately examined through behavioral or psychological testing, and will be discussed further in the next section. In summary, however, the scientific understanding of human neurological processing of receptive prosody has expanded considerably in the last two decades.

In response to this increased understanding of the neural processing of prosody, Friederici & Alter (2004) proposed a Dual-Pathway model of language processing as an expanded version of Friederici's (2002) language processing model. This Dual-Pathway model describes how semantic, syntactic, and segmental aspects of language are largely processed in the left hemisphere, while prosodic, suprasegmental elements are processed through a separate pathway in the right hemisphere. Both of these language pathways rely heavily on fronto-temporal circuits in their respective hemispheres due to "stimulus functions and processing demands." (ibid., 2004) This model of a dynamic dual pathway interaction (and a pathologically driven breakdown of it) will be crucial for our understanding of language, and in particular, prosodic processing in ASDs.

1.1.3 Prosody in ASD

Despite the surge in interest in the study of prosody in healthy individuals, scientific understanding of prosodic processing in individuals with autism is still limited. While there seems to be a general consensus regarding the impairment of prosody, the specifics of how and why are not yet well understood. In fact, as suggested by the DSM-IV diagnostic criteria, not all autistics even have problems with prosody. Paul, et. al., (2005) point out that two studies from 1975 and 2004, report rates of impairment in 4 of 7, and 47% of 30 participants, respectively. Despite the heterogeneity of symptoms in these populations, if approximately 50% of individuals with autism suffer from impaired prosody, that still translates to 1/220 of the population at large. The authors go on to say that "[w]hen such behaviors are present, however, the prosody characteristics of a person with autism constitute one of the most significant obstacles to his or her social integration and vocational acceptance. Prosodic differences are persistent and show little change over time, even when other aspects of language improve." (p. 205, Paul, et. al., 2005) Whether prosodic impairments can be improved through training remains unclear, but given the consequences of the impaired ability, clarifying the nature of the impairment and developing therapeutic strategies are well worth the endeavor.

In an attempt to specify the deficits in autistic prosody, there have been a number of studies relating to perception of sentence type as differentiated by intonation, as well as studies relating to the perception of emotional prosody. One study, for example, related to processing of types of speech acts that demonstrate the inability of autistic individuals to discriminate between interrogative-type and declarative-type prosody. (Peppé, et. al., 2007) This study also examined the perception of 'turn-end', affect, and focus as specified by pragmatic, emotional, and semantic dimensions of prosody. The results show a clear difference between the control (neuro-typical) and ASD groups, with every autistic participant showing some difficulty in producing and perceiving prosody. (ibid., 2007) The authors conclude that while receptive and expressive abilities are different, they are related, and that imitation of receptive prosody (possibly activating the Mirror Neuron network) could potentially be used therapeutically to improve expressive language. Also, the authors suggest that since the results correlate mental and verbal age, there is a pattern of "delay rather than deviance". (p. 1024, ibid., 2007) Since equivalent adult mental and verbal age is in many cases never reached in ASDs, this argument may not be valid. Despite this shortcoming, the article does contribute a good deal to the understanding of the specific range of abilities in both expressive and receptive language processing in autistics.

In terms of emotional prosody, studies have been less conclusive than studies of prosody in other domains. One hypothesis, based on Theory of Mind predictions, suggests that the general difficulty with empathic response found in people with ASDs could be responsible for emotional prosody impairments as well. Early research on this topic found conflicting results: some studies found no correlation between ASDs and difficulty in face, voice, or face to voice matching emotional paradigms, while other studies found the opposite. (see Golan, et. al., 2007 for a review) After conducting a revised study based on Rutherford, et. al. (2002), researchers found what appear to be replicable results showing that people suffering from both autism and Asperger's are indeed less capable of recognizing emotional prosody in speech. (Golan, et. al., 2007) The production of affective prosody, however, has been consistently shown to be impaired in people with ASDs. (Paul, et. al., 2005; Caria, et. al., 2011) These results support previously mentioned research on autistic prosodic processing abilities as well as, to a certain extent, the Theory of Mind hypothesis. (but, again, see Chevallier, et. al., 2011) Altogether, despite initially inconsistent findings, the more recent picture suggests that people with ASD have major difficulties with emotional/affective prosody.

Pitch Discrimination and Local Bias

Prior to asserting a model which analyzes the prosodic deficits in autistics, it is important to establish that low-level pitch discrimination abilities are intact. If a simple pitch discrimination task were impossible for a person with autism, then the question as to the nature of the prosodic impairment would be essentially solved. This is not, however, the case. Bonnel et. al. (2003) showed that people with High Functioning Autism (HFA) performed better than typically developing controls in pitch discrimination tasks. This study extended the results of previous studies showing that autistics perform better on tasks requiring low-level perceptual discrimination abilities in the visual-domain, and supports the hypothesis of an over-abundance of short-range connections within individual brain areas, such as Primary Auditory Cortex (PAC), and the visual cortex in the occipital lobe.

Russo, et. al., (2008) however, found deficient brain-stem encoding of pitch in people with ASD. These results, while seemingly contradicting other findings, have been interpreted as simply implying that abnormal pitch processing in the cortex may have a sub-cortical basis. (Hesling, et. al., 2010) Auditory discrimination ability is generally found to be enhanced in autistics, whether the stimulus is a speech or a non-speech sound. Lepistö, and colleagues (2005; 2008) demonstrated that low-level discrimination abilities applied to aspects of language such as phoneme discrimination. The HFA participants were found to have an impaired ability to extract the ‘invariant’ features (Underlying Representations) of the phonemes, however, such as when external noise obscured the phoneme being presented. The authors argue that this specificity of perception at the cost of generalization extends the Local Bias model of Happé & Frith. (1994; 2006)

The Local Bias model argues that ‘weak-central coherence’, or an impairment in the ability to consolidate information at the global level, is caused by a bias toward local information in perceptual processing, or piece-meal details of stimuli, as opposed to an actual deficit in global processing (i.e. integration of the whole). (Frith & Happé, 1994; 2006) In the case of phoneme (or invariant feature) extraction, the local bias seems to outweigh the specification of a consistent underlying representation. These findings provide support to both the earlier

Weak-Central Coherence model and the revised Local Bias model. Both models are used to explain the psychological or cognitive manifestations of ASDs, and are largely compatible with the later and more physiologically oriented theory of under-connectivity.

More recently, research has found that autistic individuals tend to be not only impaired in their spoken/auditory language consolidation modality, but also in their visual/gestural modality for language input. (Silverman, et. al., 2010) The researchers presented language stimuli in the auditory domain while showing matching videos of people gesturing. The researchers measured eye-movements to determine if a fixation on the gesturally relevant area (consistent with the theory of multi-modal gestural processing) resulted in enhanced linguistic processing ability in neuro-typical and ASD participants. They found that while consolidating visual and auditory information aided the neuro-typical participants, it actually seemed to hinder the response time for the ASD participants. The researchers controlled for mono-modal language and gestural ability and posited that the consolidation of the two respective modalities is where the difficulty lies for ASD individuals. These conclusions provide further evidence in support of Weak-Central Coherence and/or under-connectivity manifesting in ASD, both for visual and auditory language, suggesting that a consolidation, coherence, or connectivity problem is at the heart of linguistic difficulties in ASD.

1.1.4 Music and Language

Considering the difficulties that autistics have with prosody in general, and emotional prosody in particular, the natural next question would be to investigate impairments in musical abilities. Prosody is often referred to as the music of language or the melody of speech, and music is frequently described as being a language of its own. Further, music has been described as the “language of emotion”, (Levitin, 2008) or that it “sounds like emotion feels”, (Pratt, 1952) and within the Excellence Cluster ‘Languages of Emotion’ at the Freie Universität Berlin, music psychology is a major focus.

One of the most interesting and important theories in recent years regarding music and language was advanced by Steven Mithen in his book *The Singing Neanderthals.* (2006) The theory is that music is more than just auditory ‘cheesecake’, as Stephen Pinker controversially claims in How the Mind Works (Pinker, 1997; for a rebuttal, Fodor, 2000), but that it serves many evolutionary functions, such as aiding group cohesion, and indexing the emotional state of the ‘speaker’. In fact, the theory that Mithen advances is that music and language evolve from a common behavioral ancestor, a “musical proto-language”. (2006) This primitive form of communication could not be thought of as music or as a language in the modern sense, but as more of a hummed system of generally iconic and static utterances. In other words, this communicative system is similar to modern prosody, without the relevance for syntactic parsing that is one of its functions today. (see Peretz, 2006, for a similar theory)

Music has, in the last few decades, been shown to have more commonalities with language than was previously supposed. First, Lerdahl and Jackendoff developed a theory elaborating on ideas from Schenker, Bernstein and others of the generative syntax of music which proposed that harmony and meter are two governing systems of structural grammaticality in music. (Lerdahl & Jackendoff, 1977; Lerdahl & Jackendoff, 1983) This system, known as the Generative Theory of Tonal Music (GTTM) established a phrase structure of heads and specifiers, and described rhythm, or metrical structure, formally, but did not write generative rules to allow for the formation of new pieces of music. It also provided the first use of a tree branching system for describing this grammar following linguistic conventions for expressing syntactic organization in formal linguistic theory. This theory was later expanded by Martin Rohrmeier (2011) who established a system called the Generative Syntax Model, or GSM, which gave a precise set of re-write rules for phrase structure building and harmonic development, including the expansion of the GTTM’s descriptive ability for modulation, recursion, and complexity. (see Koelsch, 2012, for a brief history)

During these past decades, several neurological and behavioral studies have also been conducted on the perception and processing of musical syntax. (Patel, 2003 for an early review; Janata, 1995; Koelsch, et. al., 2005; Garza Villareal, et. al., 2011; Sammler, et. al., 2012) Koelsch, et. al., (2000; 2003; 2005) found that when contextually inappropriate chords (Neapolitan Sixth, in particular) were played, a negative ERP was elicited in the right hemisphere in an early time window. This ERP, known as the Early Right Anterior Negativity (ERAN) is generally thought to correspond to syntactic violations in music. It provides a parallel to agreement violations in language, such as incongruous gender declension, which the same study showed elicits an early going left anterior negativity (ELAN). (ibid., 2005) Further, Koelsch et al., (2013) showed that violations of hierarchical syntactic processing in music, including long-distance, non-local dependencies predicted by the GSM, result in a similar ERAN response. Neurological correlates to established syntactic and music theory provide evidence for the shared evolutionary history of language and music, while raising questions as to the neurological processing resources shared between the two.

Music has been shown, for example, to be capable of communicating meaning. (Poulin-Charronnat, et. al., 2005) In another study by Koelsch and colleagues at the Max Planck Institut in Leipzig, it was shown that music is capable of eliciting an N400, which, as mentioned previously, has long since been established as a neurological response to semantically incongruous sentences. (Koelsch, et. al., 2004) The study used a priming paradigm for congruous and incongruous adjectives in relation to the piece of music just played. One example of a congruous prime was the presentation of the word ‘wide’ following a segment of music from Strauss’s Opus 64 (Salome). When the same word was presented following the incongruous prime to an accordion piece by Valpola, an N400 was elicited. (ibid., 2004) While this particular example displays a fairly iconic meaning, other examples in the study exhibited more abstract and indexical types of semantic priming. This study shows that in a normal population, similar neural effects can be observed in the semantic priming of music or language.

Following the work of Steinhauer, et. al. (1999) described earlier, several studies have demonstrated the presence of a similar Closure Positive Shift (CPS) like response to phrase boundaries in music. (Knösche, et. al., 2005; Neuhaus, et. al., 2006; Nan, et. al., 2009) The authors of these studies noted (following Riemann's (1900) observation) that the characteristics of musical phrase boundaries are similar to those of language. In music, phrase boundaries are often marked by a phrase-final note lengthening, followed by a pause. In the first study, by Knösche, et. al., the researchers observed this effect, as predicted, in response to musical phrase boundaries, similar to the response elicited by prosodic boundaries. (2005) They called this ERP the 'music CPS'. This ERP component is characterized by a different time course than that found in response to language, however. The 'music CPS' was found following the onset of the first post-phrase-boundary note as opposed to occurring during the boundary, or pause itself. The latency was much later than that of the linguistic study, yet had a similar peak of around 500 ms, with a centro-parietal maximum. (Knösche, et. al., 2005) The authors attribute this latency difference to earlier cues in the linguistic stream, such as a varying prosodic contour. Later, Neuhaus et al., (2006) found effects of harmonic closure type, by comparing phrases that ended on a dominant (fifth scale degree, half-cadence) vs. tonic (first scale degree, full-cadence) note. They found a more anterior distribution of the music CPS in response to these half cadence conditions, perhaps suggesting that some effect of expectation contributed to this component when the musical piece remained unfinished. Additionally, a more recent study (Silva, et al., 2014) has found a more similar ERP component occurring during the phrase boundary. This ERP is a positive going potential in response to the phrase offset, with a maximality at electrode Cz for well-formed musical phrases. Whichever of the two phrase-boundary response components represents the musical homologue of the language CPS, the similarity of the ERPs provides a good methodology for contrasting the impact of prosody and music in signaling the structuring processes responsible for the division of phrases.

Other parallels between music and language have shown a the correlation between being a native speaker of a tonal language and having perfect pitch. (Deutsch, et. al., 2009) The researchers found that when comparing the prevalence of perfect pitch in native English speakers and Native Chinese speakers in a Chicago music conservatory, the percentage of native Mandarin speakers with perfect pitch was significantly higher than those with English as a native language. (ibid., 2009) Deutsch and colleagues also demonstrated that when a speech phrase is played repeatedly (10 times) to participants, they perceive the phrase as 'sounding like singing' more than 'sounding like speech'. (Deutsch, et. al., 2008; Deutsch, et. al., 2011) The authors classify this perceptual response as an auditory illusion, but the illusion actually begins to reflect the reality of pitch change in the utterance. Thus, the boundaries between speech and singing are not only difficult to define, but can change based on developmental or contextual factors.

Nevertheless, the intertwined nature of the two depends crucially on fundamental frequency, f(0). Languages with lexical tone mark a contrast between words that differ in f(0) height or contour, while other types of languages (e.g. pitch accent, stress-timed) mark f(0) contrast only on a phrasal level or with lexical contrastive stress. This distinction results in more refined pitch discrimination ability in native speakers of tonal languages. When phrases are perceived as being sung, this illusion is due to the accurate perception of f(0) pitch height and change, instead of the course contrastive relative pitches typically perceived at the phrasal or sentence level. These studies document the interaction of music and language through the common focal point of pitch and suggest that a more refined perceptual ability for pitch would allow a person to have heightened linguistic and musical abilities. (Deutsch, et. al., 2008; Deutsch, et. al., 2011) As the studies of people with ASDs previously mentioned, however, improved perceptual abilities do not always translate into broader linguistic facility.

Part of the difference in auditory processing regarding music and language relates to the contrast of information with temporal and spectral complexity. Zatorre, et. al., (2002) found that the function of the two cerebral hemispheres is divided into specialization for temporal complexity (left) and spectral complexity (right), which corresponds to the physical realities of language and music respectively: "speech is highly dependent on rapidly changing broadband sounds, whereas tonal patterns tend to be slower, although small and precise changes in frequency are important. We argue that the auditory cortices in the two hemispheres are relatively specialized, such that temporal resolution is better in left auditory cortical areas and spectral resolution is better in right auditory cortical areas." (p.37, Zatorre, et. al., 2002) In contrast to the general understanding of language and music overlap, this study implies that there is a functional division between the hemispheres dedicated to the processing of each domain. (Aasland, et. al., 2003) Studies have also shown that the fundamental frequency and prosodic information of language is largely processed in the right hemisphere, with the STG playing a major role. (Friederici & Alter, 2004; Zhang, et. al., 2010; Sammler, et. al., 2010) This suggests that the global pitch-dependent features of prosody may lead it to be processed in more of a frequency specialized, or spectrally dedicated, right hemisphere network assembly.

Music and language are generally considered to be processed using shared processing resources, particularly for syntax, (Patel, 2003; Patel 2005) but also for semantic content. (Koelsch, et al., 2004) Many right hemisphere structures are homologous to those left hemisphere areas dedicated to language processing, yet both language and music rely on both hemispheres to coordinate perceptual and organizational operations. Hierarchical harmonic organization of music relies heavily on left-hemisphere structures, while the initial stages of prosodic processing in language are largely confined to the right hemisphere pathway.

1.1.5 Music & Autism

The most commonly known examples of autistic musical ability likely relate to savant syndrome. Awareness of savant syndrome by the scientific community can be traced back to at least 1751, with the prodigious mathematical calculation ability of Jedediah Buxton. (Foerstl, 1989; Treffert, 1988) This syndrome, despite its prevalence in the public imagination, occurs infrequently, with many estimates around .06% of mentally disabled people in institutions. (Treffert, 1988) Even in autism, the most common disability with which it occurs, the condition is rare, with a prevalence of roughly 9.8%. (Rimland, 1978) Despite the rarity of savants, there are particular skills to which the condition tends to be limited, particularly calendrical calculations, multiplication and division, memorization, art, and music. Within all these categories of skills, there is a distinction between talented and prodigious savants. Those who have abilities not expected with their hindered development are known as talented savants, while those who have abilities surpassing the abilities of typically developed people or specialists, prodigious savants. (Treffert, 1988) While there is a technical distinction between these two groups, both have been found to use rule-based approaches for their abilities. (O'Connor & Hermelin, 1984)

The existence of autistic musical savants suggests that simply having the diagnosis of ASD does not mean that musical abilities are impaired. On the contrary, it appears that musical abilities can even be heightened. The recognition that this is based on rule-based calculations instead of rote-memory (although excellent memory is also a common feature), shows that these savants have the ability to reconcile syntax with music. While this is a common ability found in neuro-typicals, it suggests a reliance on inter-hemispheric coordination of neural processing, due to hierarchical structure building operations taking place mostly in the left hemisphere and with spectral complexity representations being found in the right. In fact, right hemispheric specialization for music processing may prove advantageous to individuals with ASDs. As seen earlier, reversed hemispheric asymmetry is well documented in ASD participant pools, particularly for the IFG. (De Fosse, et al., 2004; Hodge, et al., 2010) The larger right hemisphere IFG, which has been shown to be highly involved in music processing, (Sammler, et al., 2012) may not only preserve musical abilities in ASDs, but may indeed allow for superior musical abilities, and perhaps even the type of specialization seen in musical savants. The impairment seen in receptive and expressive prosody in autism creates challenges for these interpretations, however. The well established correspondences between music and prosody would suggest that musical ability should be impaired in people with autism as well. Since, as we have seen, this is not the case, there seems to be a dissociation between the two abilities. Why this dissociation exists is not well understood.

It has been shown, however, that in certain limited situations, autistic individuals have trouble processing temporal complexity, suggesting that the difficulties autistics have with language processing could be due to a problem interpreting its temporal complexity. (Samson, et. al., 2011) Perhaps the rapid, large pitch shifts which characterize prosody require the incorporation of both detailed temporal and spectral processing architecture.

Music Therapy

Music as a therapeutic device for ASD has received substantial attention. (See Simpson & Keen, 2011; and Accordino, et. al., 2007, for reviews) Several different strategies with a variety of goals have been implemented, most using music listening in various situations, with the goals of improving socialization, communication, or behavior. In general, there have been positive results of music used as a therapeutic device in children with autism, not only for improved socialization and behavior, but also for communication. In one study, researchers found that when words or gestures were put to music, they were much more likely to be remembered than words just put to the same rhythm. (Buday, 1995) This test shows that the widely recognized mnemonic characteristics of musical association apply to people with ASD, even if a follow up study assessing their use of these words in a broader social environment has yet to take place. (Simpson & Keen, 2011)

Another study found that the use of improvisational music therapy can be beneficial in improving the communicative abilities of those suffering from autism, while simply listening to music saw a reversal of the positive effects. (Edgerton, 1994) This finding suggests that when those undergoing this treatment produce music themselves, they are engaging in expressive communication in a musical domain, and through engaging this domain, they also have an improved response to communication. These results parallel those found by Peppé et. al. (2007) regarding the interaction of expressive and receptive abilities in language mentioned earlier. If music can improve communicative ability while the 'music of language' remains impaired, where is the disconnect?

Melodic Intonation Therapy is another therapeutic technique for populations with limited language ability. (Norton, et. al., 2009) Unfortunately, this therapy has, for the most part, only been applied on a large scale to patients with aphasia. It will be interesting to see whether this method is successfully applied to people with autism. The therapy entails slowly repeating 'sung' sentences or phrases to a person with disordered language in order to use an intact ability (singing) to modify, or help the relearning of a damaged ability (speaking). (ibid., 2009) The technique, as it was originally designed, used only two alternating tones to help reteach aphasics the intonational patterns of spoken language. Considering that imitation, pitch perception, and rule-based grammar are all intact abilities for most individuals with autism, it would be natural to suppose that this type of therapy could also be beneficial.

Music & Language in ASDs

There has been a scarcity of research investigating what the differences in processing abilities are for language and music in ASD. One study, initially following the functional under-connectivity hypothesis found that while there seems to be a functional under-connectivity for speech processing in autistics, the left hemisphere language circuits (such as Broca's area) demonstrated increased activation for music. (Lai, et. al., 2012) The authors suggest that the reasons for this increased activation are not due to functional under-connectivity as a whole, but under-connectivity in the dorsal language processing stream of the left hemisphere, (Hickok & Poeppel, 2004) known for its role in "high-level" (p. 972, Lai, et. al., 2012) language processing. According to Hickok & Poeppel's (2004) model, this dorsal stream (which passes through the sensory and motor cortices) is responsible for maintaining parity between acoustic/articulatory functions, in a similar manner to the way the MN system operates. This explanation fails to address the central issue of autistic language impairment, however, which is the left/right hemisphere coordination involving prosody. While this study proposes concrete specification of the nature of under-connectivity in autism, the theory remains unproven.

The authors take an additional approach to explain the deficit in language ability, namely that it is due to impairments in attentional abilities: "[o]ne possibility for discrepancies between music and language functions in autism and models that propose long-range disconnection may be a speech-specific (and in general, social-information-specific) attentional deficit." (p. 973, 2012) This social attentional deficit does seem to be at the core of autism in general. (Baron-Cohen, 1995) The under-connectivity hypothesis predicts such a deficit due to evidence of an impairment in the connectivity between frontal and parietal lobes in autism, (Just, et. al., 2012) networks known to be involved in attention and orienting. (e.g. Corbetta & Schulman, 2002) Autistic individuals frequently lack an interest in the initiation of joint attention (IJA), which has been shown to be a significant predictor of language development. (Mundy & Gomes, 1998) The IJA requires the close collaboration of temporo-parietal and frontal regions, (Just, et. al., 2012) due to the cognitive control necessary to voluntarily orient attention. Parieto-frontal systems have been repeatedly demonstrated to be involved in attentional orienting in several domains, both for voluntary, or goal-directed orienting, as well as involuntary orienting to relevant stimuli. (Corbetta, et. al., 1995; Corbetta & Schulman, 2002) The deficit in these long-range connections in ASD may result in impairments to elements in this parieto-frontal attention network, albeit with preserved attentional functioning in the parietal lobe itself. (Just, et. al., 2012) With broad deficits in the attentional network, it is unsurprising that autistic individuals demonstrate reduced interest in social-information as a whole, much of which is dependent on joint attention. Additionally, these deficits may affect prosody, and thus, language more broadly, as the purpose of prosody is often to draw attention to some element of the speech stream.

1.2 The Present Study

Based on the Dual-Pathway model of language processing outlined earlier (Friederici & Alter, 2004), it appears that for many people with autism, something goes wrong with the processing of one of the two language pathways. The individual pitches, combinatory syntax, and semantic content can all be perceived on their own, yet the difficulty lies in combining the broader acoustic features of the speech signal. Whether to aid in segmenting, or combining elements of the grammatical form, or cue the listener to the emotional state of the speaker, together, these fail to create a single coherent global whole. The theories of under-connectivity and weak-central coherence provide a framework through which to investigate this problem.

Further, music is capable of being processed syntactically by specialist savants and many others with autism. It is also capable of improving communicative ability, something which both music and prosody can do for typically developing individuals. It appears that here is where the differences between music and prosody are most manifest. Why is it that music can be an effective communicative aid to autistics while prosody cannot? Is it that within music, individuals with

ASDs are able to use global melodic features to guide syntactic operations yet cannot with prosody? Or is it that the function of melody in music is not generally to draw attention to particular hierarchical constructions within the music, but that the rhythmic and harmonic structure serves to support the melody?

The differences between melody and prosody represent the larger differences between music and language on a fundamental level. While the two share processing resources, similar hierarchical constructions, and the ability to communicate meaning, the dissociation between music and language in autism is crucial to understanding the natures of music and language. Similarly, understanding the differences between music and language, in structure, function, and neural processing will be crucial to understanding the communicative disability in the disorder of autism. Since prosody is the melody of language, yet in this instance initiates such a different response from that to music, comparing responses to the two domains in neuro-typical, and ASD participants will be informative for further specifying where these differences lie.

1.2.1 Phrase Boundary Processing

The Weak-Central Coherence (WCC) and Local Bias models of autism described earlier (Frith & Happé, 1994; Happé & Frith, 2006), propose that autism creates a processing bias toward the local information, or fine grained detail, at the expense of the global whole. This local bias manifests itself in instances such as enhanced pitch and phonemic discrimination ability, with difficulty consolidating supra-segmental information into larger coherent global form in the context of language. Taking this a step farther, individuals with autism have little difficulty perceiving the local elements of syntax and prosody, but are impaired in their ability to attend to the global aspects of the two and consolidate the two language processing pathways. If individuals with this type of disorder have little difficulty consolidating larger, global aspects of music, however, this would serve as counter evidence to strong forms of the WCC, Local Bias, and under-connectivity hypotheses, especially since music and prosody share extensive processing resources.

The syntax of music is beginning to be fairly well understood. As mentioned earlier, harmonic functions serve as the grammatical categories of musical phrases, and (at least for Western Harmony) a generative framework is well underway. Further, similar types of acoustic features mark phrase boundaries in music as in language. In both domains, there is frequently a pre-boundary lengthening of a syllable or note and the phrase boundaries are frequently marked by pauses in both music and language. The neurological responses to these features of phrase boundaries are also comparatively well understood. As mentioned earlier, both types of phrase boundaries elicit a CPS, or a Closure Positive Shift, a positivity occurring soon after the phrase boundary and lasting between 300-500 ms.

The CPS remains to be investigated in an ASD population. Following Steinhauer et. al., (1999), the CPS is a prosodically driven electrophysiological component that indexes syntactic boundary processing. It would then be natural to hypothesize that if autistic individuals have difficulty interpreting prosodic information to cue syntactic information in language, a CPS would have a different manifestation at phrase boundaries. Following the evidence that individuals with autism do not exhibit the same difficulty incorporating musical information into a cohesive form, then we should expect that musical phrase boundaries will elicit a CPS in this group. It remains possible that neither the musical nor linguistic CPS will be elicited, but considering the tendency toward accurate and preserved musical perception in ASDs, this seems unlikely.

1.3 Hypotheses

Neuro-typical participants will exhibit a centro-parietally maximal CPS at phrase boundaries in language occurring almost immediately after the onset of the phrase-boundary, while ASD participants will demonstrate either no CPS, or a CPS with markedly different distribution, amplitude, or latency.

Neuro-typical participants will exhibit Garden Path N400/P600 effects in response to the spliced incongruous prosody/syntax condition, while ASD participants will not.

Both neuro-typical and ASD participants will exhibit a music CPS at phrase boundaries in music to the Phrased and Phrased No-Pause conditions, and neither group will exhibit a musical CPS at phrase boundaries in music in the UnPhrased condition.

Both groups will show an effect of Cadence, or harmonic closure, on the music CPS.

Both groups will show a positive electrophysiological response during the musical phrase boundary.

2 Methods

2.1 Participants

For this study a total of 34 participants were tested. Thirteen (3 women) had a diagnosis on the Autism Spectrum of High-Functioning Autism, Asperger's Syndrome, or both from a licensed physician using the "gold standard" of diagnosis, the Autism Diagnostic Observation Schedule (ADOS; Lord, et al., 2000), and Autism Diagnostic Interview - Revised (ADI-R; Lord, et al., 1994), or the internationally recognized criteria of ICD-10 (WHO, 1992) and/or the DSM-IV TR (APA, 2000). These participants were recruited from the Berlin/Brandenburg metropolitan area either through psychiatrist recommendation, or through advertisements posted on internet forums or at organizations related to autism. The ages of the ASD participants ranged from 23 to 54 with a mean age of 36.58 (SD 9.55). Of the ASD participants, one was rejected for excess artifacts in the EEG recording, leaving a total of 12 for analysis.

Twenty one (7 women) control participants were tested. These participants were recruited from flyers posted around the Freie Universität Berlin. All participants gave informed consent and were paid for their participation. Ethics approval was obtained from both the FU Berlin, and Tulane University. Of the neuro-typical participants, one requested to withdraw from the study, and two others were rejected due to excess artifacts, leaving a total of 18 for analysis. The ages of the neuro-typical participants ranged from 21 to 56 with a mean age of 28.22 (SD 8.077). The difference in the ages of the two groups was significant (t = 2.5, $p = 0.021$), but when analyzing an age-matched subset of the neuro-typical group's behavioral and EEG data, the results were virtually identical to the larger group, so for purposes of clarity, only the larger group's data will be reported.

Exclusion criteria included a history of neurological problems (outside of ASD), a family member with an ASD (for the neuro-typical group), or alcohol or drug dependence, which no participants reported having.

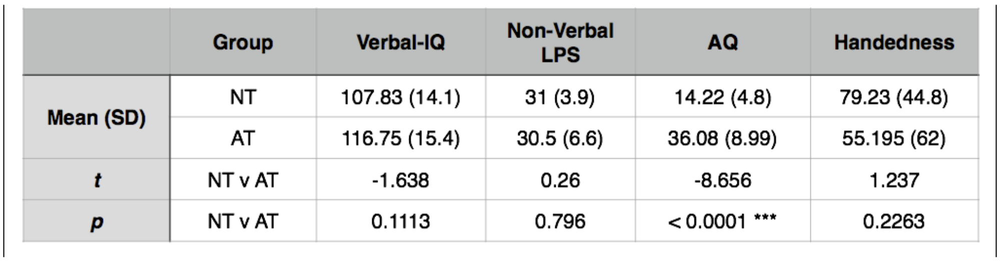
Neuropsychological Battery

In addition to information about age, and other basic demographic information, participants were all given a series of questionnaires, including the German translation of the Autism Questionnaire (Baron-Cohen, 2001), the Mehrfachwahl Wortschatz-B (MWB) test for Verbal IQ (Lehrl, 1995), the Leistungprüfsystem (LPS), Section 4 test for Non-Verbal IQ (Horn, 1965), a Handedness Questionnaire (an adapted version of the German translation of Oldfield, 1971), and a questionnaire designed to determine the musical background of participants.

The Autism-Spectrum Quotient (AQ) is a screening test designed to test "the degree to which an adult with normal intelligence has the traits associated with the autistic spectrum." (p. 5, Baron-Cohen, et al., 2001) It is not intended as a diagnostic strategy, but includes questions designed to asses perceptual differences, as well as preferences regarding social interaction and repetitive behavior. Generally, a cutoff of score of >32 is used to classify those individuals with a high degree of ASD related traits. None of the neuro-typical participants scored above 23.

The MWB is a vocabulary based verbal IQ test capable of being administered in a matter of minutes. It consists of 32 questions containing a single real German word, and several phonologically plausible non-words, from which participants must choose the real word. The LPS is a multi-dimensional test designed to examine a wide range of reasoning tasks. Section 4 of this test is a logical, non-verbal reasoning test in which the task is to identify patterns and outliers. Only this section was chosen as it is a standard test of non-verbal reasoning ability capable of being administered in a short amount of time. Only raw scores from this section are shown since an IQ as calculated from this section alone would not be representative. Two participants in each group had musical training, and the results of the other questionnaires are shown in Table 1.

Table 1: Results of Neuropsychological Battery



	Group	**Verbal-IQ**	**Non-Verbal LPS**	**AQ**	**Handedness**
Mean (SD)	NT	107.83 (14.1)	31 (3.9)	14.22 (4.8)	79.23 (44.8)
	AT	116.75 (15.4)	30.5 (6.6)	36.08 (8.99)	55.195 (62)
t	NT v AT	-1.638	0.26	-8.656	1.237
p	NT v AT	0.1113	0.796	< 0.0001 ***	0.2263

While the two groups are not significantly different in terms of handedness, a high degree of left-handedness and ambidextrousness was found in the ASD group, necessitating recruitment of left handed and ambidextrous participants in the NT group. The prevalence of left and mixed-handedness in ASDs is well documented, and the associated neuroanatomical asymmetry may be related to the cognitive impairments typified in ASDs. (Soper, et al., 1986; Floris et al., 2013)

2.2 Stimuli

2.2.1 Language Stimuli

The language stimuli were adapted from the stimuli used by Pannekamp, et al., (2005). These stimuli, which were initially directed toward young children, were adapted slightly to make the task appropriate for older participants. The speech samples were recorded at the Center for General Linguistics (ZAS) in Berlin, Germany. These were recorded to DAT on a TASCAM DA-20 MKII, (TASCAM, Montebello, USA) and a Sony DTC-690, (Sony, Tokyo, Japan) with a Sennheiser (Sennheiser Electronic GmbH, Hanover, Germany) Condenser MkH 20 P48 O (omni), run through a Behringer Ultragain Mic2000, (Behringer, Wellich, Germany) and then exported as .aiff with Adobe Audition 1.5. (Adobe Systems, San Jose, USA) The speech samples were spoken by a native German speaking woman with knowledge of the experiment and the phrase boundaries that were being examined. These were then uploaded onto Audacity 2.0.3 (Mazzoni, http://audacity.sourceforge.net) and edited such that there were between 150 and 200 ms of silence preceding and following each speech segment. Then, breaths and other background noises were carefully silenced to reduce possible confounding features of pauses that were not expressly the subject of the experiment. The samples were in Mono, but were doubled to stereo and then Normalized to max -1 dB, and DC offset was removed.

The sentences were grammatically all of the same form which is as follows:

A. ‘Kevin verspricht Sophie zu schlafen, und ganz lange lieb zu sein.’
Kevin promises Sophie to sleep, and to be a good boy for a while.

B. ‘Kevin verspricht, Sophie zu küssen, und ganz lange lieb zu sein.’
Kevin promises to kiss Sophie, and to be a good boy for a while.

These sentences differ in the argument structure of the two verbs. In the above examples, the verb ‘promises’ takes no accusative argument in condition B, while it takes ‘Anna’ as its second argument in condition A. Also in the above examples, the second verb ‘helfen’ (to help) in condition B takes ‘Anna’ as its second argument, while ‘schlafen’ (to sleep) is intransitive. The sentences follow the same word order, yet the German intonation patterns with which these two sentences are spoken differ, providing a perceptual distinction between the types of structures being produced, aiding in the hearer’s parsing of the sentences. The most perceptually salient distinctions between the two sentences occur at the Intonational Phrase (IPh) boundaries, which correspond in the above examples to the commas, and in the auditory versions, to pauses. This means that in the condition A, there are two phrase boundaries resulting in three IPhs, while in the intransitive condition, there is only one phrase boundary resulting in two IPhs.

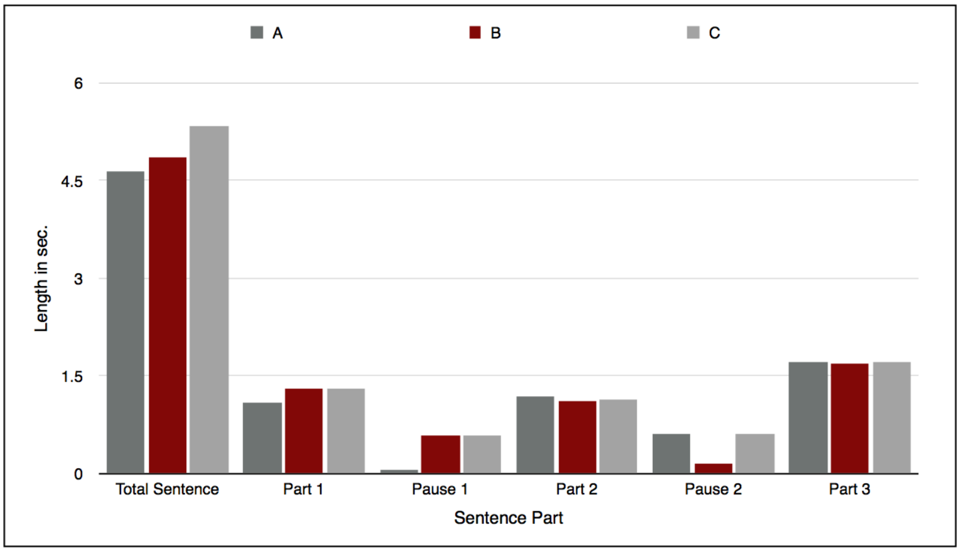
Pauses at the first IPh in condition B (and thus also C, described below) were edited to a consistent length of 600 ms. This was done in an attempt to make the speech samples more consistent, both internally and with the music stimuli. Many of the pauses were already at this length and previous studies (i.e. Pannekamp et al., 2005) have found that this pause length is typical for these types of sentences. For the same reasons, the pauses in condition B were edited such that the first pause was 600 ms, while the second pause was consistently under 200 ms.

The two types of sentences were then carefully cross-spliced during the alveolar-dental closure at the beginning of the affricate /ts/ in the word ‘zu’, so that the sentence up until that point was drawn from condition B, while the sentence following that point was acoustically identical condition A. This resulted in sentences as follows:

C. *Kevin verspricht, Sophie zu schlafen, und ganz lange lieb zu sein.
**Kevin promises to sleep Sophie, and to be a good boy for a while.*

These sentences are prosodically and syntactically incongruous, and following linguistic convention, are marked with an asterisk. In total, 48 sentences of each of the three conditions were produced resulting in a total of 144 sentences. These are listed in Appendix I. Each recorded sentence is approximately 4 to 5 seconds in length. Figure 1 shows an exact breakdown of lengths of each sentence part, where Part 1 is the length of the sentence up to the first pause, Part 2 is the length of the phrase between pauses, and Part 3 is the final phrase of the sentence. Part 1 is shorter for condition A, since the final syllable in the other two conditions is lengthened to signal a phrase-boundary. Also, the first pause in condition A is almost non-existent as there is no phrase-boundary. All conditions differ slightly in Part 2, since condition C is spliced from the two other conditions. Part 2 is partially comprised of the second verb of the sentences, the onset of which serves as a trigger for measuring Garden Path effects. These second verbs are on average 573 ms (70 ms SD), suggesting that a CPS elicited by the phrase boundary following this verb (i.e. at the onset of Pause 2) is almost certain to interact with a P600 triggered from this verb’s incongruous prosodically cued argument structure.

Figure 1: Language Stimuli Average Lengths



2.2.2 Music Stimuli

The music stimuli were originally composed for this project following basic conventions of western musical form. The samples were composed in Sibelius First (Version 6.1.5, Avid Technology, Burlington, USA) monophonically with a realistic acoustic piano midi sound. All musical compositions were exported at 16 bit 44.1 Khz and edited and normalized using Audacity 2.0.3. (Mazzoni, http://audacity.sourceforge.net) The total length of each track equals 40.5 seconds, which includes approximately 600 ms of silence prior to and following each track, and 16 bars of music at 100 bpm with an anacrusis. The choice of 100 bpm allowed for musical presentation at a moderate tempo, while making each quarter note consistent at 600 ms. These were then normalized to max -4 dB, in order to have a perceptually similar volume to the language stimuli, with DC offset.

Each musical sample followed the same structure, which was two four bar phrases creating an 8 bar piece, which was repeated. (see Appendix I, for an example of the musical notation) Each phrase began with an anacrusis of the same length (one beat, 100 bpm), and each phrase ended on a tone at the first (full cadence), third, or fifth scale degree (for half cadences), all suggesting a tonic chord. These notes all occurred on the first beat of a measure, which has consistently been identified as the strongest beat of the measure (see Lerdahl & Jackendoff, 1983 for a review), and the phrase final note was the longest note that had in the piece. These pieces were then reviewed by a professional composer for editing and corrections. As the music CPS has previously been identified as occurring in response to the first post-boundary note, an additional anacrusis and strong beat root note were added at the end of each piece to provide a fourth post-boundary note from which to measure EEG response. The addition of this final note created complications, however, as this final note was always a single tonic quarter note, while previous anacruses were composed of either two eighth notes (25%) or a single quarter note (75%). Thus, while counter-balanced across conditions and not affecting the analysis of music CPS, the imbalance of these note lengths compromised the ability to compare effects of cadence on post-boundary music CPS, due to the presence of more auditory evoked potentials (AEPs) in the half-cadence condition.

The final notes of each musical phrase differed between conditions in length and the presence of a rest, or notes between phrases. Previous studies have only compared conditions in which a rest is contrasted with a similar metrical period 'filled in' with notes. This comparison results in inconsistent electrophysiological responses, since the comparison of a response to a note preceded by silence and a response to a note preceded by other notes leads to markedly different AEPs, or N1/P2 components being exhibited. Thus, the three categories of musical comparison are as follows (all stimuli can be found in Appendix II):

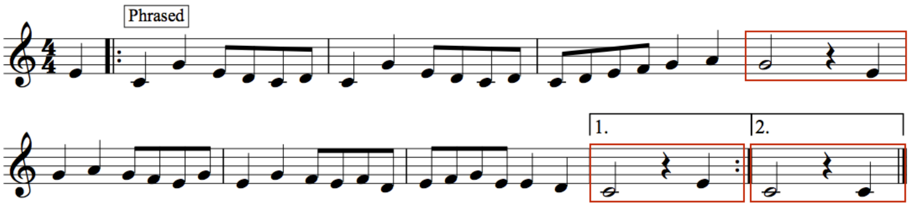
Figure 2: Phrased

Phrased (Figure 2): Musical phrase final characteristics, pre-boundary half note followed by a pause: quarter rest with a duration of 600 ms;

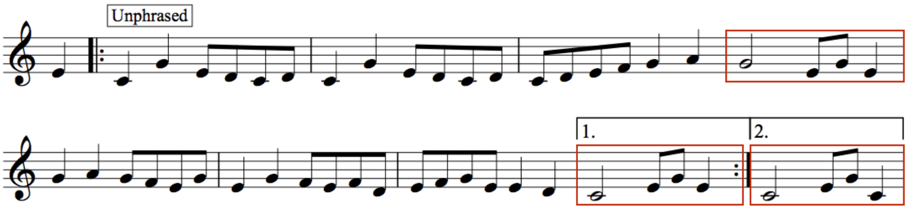
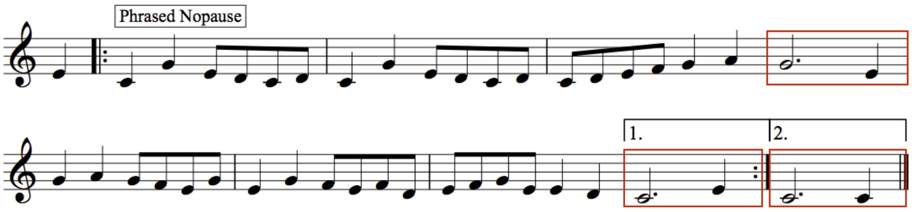
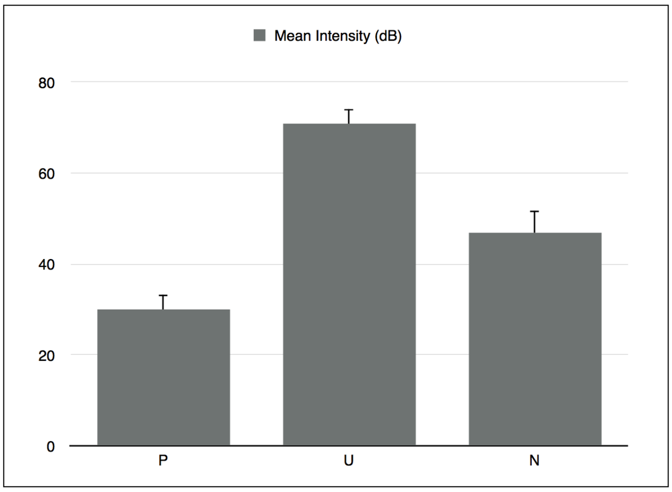
Figure 3: UnPhrased

UnPhrased (Figure 3): Musical phrase final characteristics, pre-boundary half note followed by notes: quarter rest (600 ms) filled in with either a quarter note, or two eighth notes;

Figure 4: NoPause

Phrased NoPause (Figure 4): musical phrase final characteristics, boundary dotted half note of 1800 ms duration with gradual decay.

The third condition was designed to provide evidence to determine if the musical CPS associated with musical phrase boundaries is dependent on the presence of a pause (or rest), or if other phrase final characteristics are sufficient to elicit it. The names of the first two categories are used by convention, even though as Nan et al., (2009) mention, the UnPhrased condition is really 'less phrased' since there are other phrase signaling cues in music in addition to a pause or rest, such as the presence of a longer phrase-final note, the use of strong beats, degree of harmonic closure, and the structural convention of 4-bar musical phrasing.

In order to define the differences between conditions, it was necessary to compare the acoustic characteristics of the phrase-boundary. As one condition, NoPause (N) simply decays more slowly than the Phrased condition (P), it is necessary to determine the average intensity during the time window being examined. As the phrase-boundary consisted of the 600 ms preceding the onset of the post-boundary phrase (1200 ms following the onset of the phrase-final note), the average intensity of this time window was analyzed using sound analysis software Praat (Boersma & Weeninck, 2013) for all three conditions. These values are summarized in Figure 5. Then pairwise comparisons were conducted using two-tailed t-tests using statistical analysis software R (R Foundation for Statistical Computing, Vienna, Austria) to ensure that the conditions were significantly different. All comparisons resulted in p values $<$ 2.2e-16, showing that the intensity across these periods is highly significantly different.

Figure 5: Average intensity by condition at phrase boundary: 600 ms



Previous studies examining the musical CPS have also only looked at the electrophysiological response to musical phrase boundaries in the middle of musical phrases, and have not determined if the ends of musical pieces also elicit a CPS. The presence of the repeat in the pieces seeks to look more closely into this phenomenon in order to determine if the CPS is a cognitive chunking mechanism responsible for consolidating the previous information perceived, or if it is a predictive mechanism, building hierarchical structure and making guesses about what will come next. This structure also allows for the possibility of investigating repetition effects, in order to determine whether any musical phrase-boundary processing effects change as a function of presentation novelty.

2.3 Task

The two types of stimuli were presented using Presentationő software (Neurobehavioral Systems, Berkeley, USA) in separate blocks with a duration of approximately 30 minutes each. Stimuli were presented in randomized order, and block presentation was counterbalanced among participants. Including the completion of consent forms and questionnaires, application of the EEG cap, and participation in the experimental blocks, the entire procedure lasted approximately 120 minutes. Following the presentation of each sentence the participants were asked “Wie natürlich fanden Sie den letzten Satz?” or “How natural did you find the last sentence?”. The participants then responded on a 5 point scale from “Completely Unnatural” to “Completely Natural”. The corresponding task for the music stimuli was to answer the question “Wie natürlich fanden Sie das letzte Stück Musik?” or “How natural did you find the last piece of music?”, with an identical 5 point response scale. This task’s primary purpose was to maintain participant attention during the stimulus presentation, but behavioral responses were also analyzed to determine whether ‘naturalness’ ratings varied across different conditions. Previous studies (Knösche, et al., 2005; Neuhaus et al., 2006) used a task requiring participants to identify out of key notes, which as Nan et al., (2009) point out, is much more cognitively demanding for non-musicians. Thus, while Nan, et al. (2009) found music CPS responses in non-musicians due to the use of a different task, Neuhaus et al. (2006) did not. For this reason, a task was chosen that would maintain participants’ attention while not being too demanding for those with no musical background.

2.4 EEG Recording

All EEG recordings were conducted at the Dahlem Institute for Neuroimaging of Emotion (DINE) in an electrically shielded room using a 34 Ag/AgCl electrode configuration distributed according to the international 10-20 system. While using 64 electrodes is a more standard recording strategy, the length of the study and attempting to avoid discomfort in the ASD participants encouraged the use of a quicker setup procedure. In addition to 27 scalp electrodes, (see Appendix III) 4 electro-oculogram (EOG) electrodes were attached to the outer canthi (2) of both eyes, and above and below the participants' leading eye (2) in order to aid in the identification of eye-blink, horizontal eye movement, and facial muscle movement artifacts. Additional electrodes were attached at both mastoids, of which the right mastoid was used as a reference electrode. A grounding electrode was attached to the nape of the neck. Recordings were made with Brain Vision Recording Software (Brain Products GmbH, Gilching, Germany) at a sampling rate of 500 Hz. All impedances were kept below 5 kΩ.

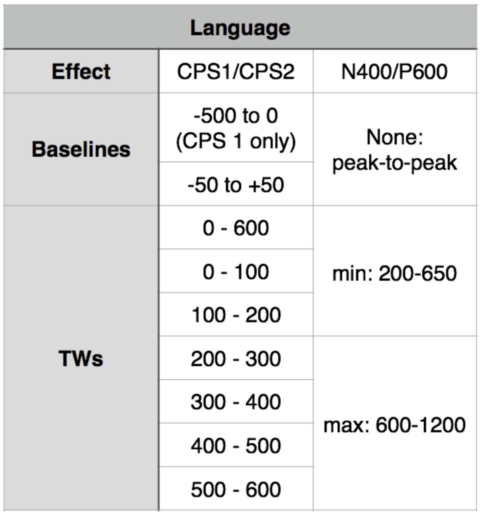
2.5 EEG Data Analysis

2.5.1 Language

EEG Recordings were analyzed using EEProbe software. (ANT Software BV, Enschede, Netherlands) Eye-blinks and facial movements were then rejected manually. Participants with more than 40% of trials rejected were not included in the analysis. A high pass filter of .5 Hz was used to remove slow drifts, and a low pass filter of 30 Hz was used to remove high-frequency oscillations and line noise.

In the language stimuli, epoch measurements were time-locked to the beginning of the pauses associated with the IPh boundaries. Previous studies have shown the CPS to have a central/parietal distribution with an almost immediate onset following the IPh. (Steinhauer, et al., 1999; Männel & Friederici, 2011) Multiple baseline correction regions were used to compare the stability of the effects observed during the different conditions. (see Table 2 for a detailed chart) Prior to the first IPh, a baseline of the 500 ms immediately preceding the IPh was used to calculate effects of phrase-boundary on electrophysiological signal. This large baseline correction window ensured that any variability across the different conditions would be accounted for in the region immediately preceding the boundary. Additionally, a baseline region of -50 to +50 ms relative to the onset of the IPh was used to confirm the results seen using the previous baseline, as well as provide a means for comparison for the later CPS using the same baseline. The fact that it includes 50 ms during which the CPS is already observable means that it eliminates some of the robustness of the observed effect. This baseline was also used for the second IPh of the sentence, as the preceding 500 ms included the variability across conditions of different phrase final verbs (and indeed sometimes the previous IPh) making the larger preceding baseline inappropriate. The pauses following the first IPh in conditions B and C were 600 ms, so the time windows (TWs) examined for this region were a large window of 600 ms, with additional comparisons for each 100 ms interval within this time window, in order to determine the nature and time-course of the CPS effect. The same TWs were used for the second IPh, which in conditions A and C was also 600 ms.

Table 2: Language TWs & Baselines



Language		
Effect	CPS1/CPS2	N400/P600
Baselines	-500 to 0 (CPS 1 only)	None: peak-to-peak
	-50 to +50	
TWs	0 - 600	min: 200-650
	0 - 100	
	100 - 200	
	200 - 300	max: 600-1200
	300 - 400	
	400 - 500	
	500 - 600	

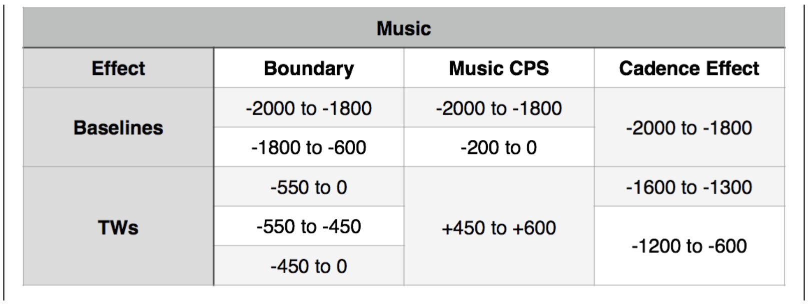
Additionally, a trigger was placed at the beginning of the second verb of the sentences in order to compare Garden Path effects in response to the prosodically incongruous spliced condition C and the other conditions. Previous studies (i.e. Steinhauer, et al., 1999) used a trigger of 200 ms following the onset of Verb 2 to examine Garden Path effects, but the placement of the trigger is somewhat arbitrary as it is uncertain when individual participants recognize the presence of an incongruity. Thus, for simplicity's sake, a trigger at the onset of the verb was chosen, with widely defined TWs (200 to 650 min; 600 to 1200 max) to examine these Garden Path effects. As Verb 2 was on average 573 ms, the TWs under examination coincide with the TW immediately preceding, and during the CPS. Additionally, these epochs were filtered with a 5 Hz low-pass filter to ensure that the peaks found were not due to alpha artifacts, but were indeed representative of the cognitive processes under examination.

2.5.2 Music

In previous music CPS studies (Knösche, et al., 2005; Neuhaus, et al., 2006; Nan, et al., 2009), epoch measurements have been time-locked to the onset of the first note following the phrase boundary. These previous studies suggest that unlike the language CPS, the music CPS does not occur during the pause between phrases, but following the first note after the phrase boundary (i.e. the first note of the new phrase). Thus, the epochs were time-locked to this note in each of the conditions. In these same previous studies, baseline references were taken from both the 200 ms preceding the previous phrase's final note, and the 200 ms preceding the first note of the post-boundary phrase. The latter baseline choice, despite being temporally closer to the time window of interest presents challenges, since in the UnPhrased condition, notes are present preceding the post-boundary note, while in the other two conditions, there is either a rest, or a decaying note, respectively. This difference leads to the presence of the onset components N1 and P2 during the baseline in one condition, and no such response in the others.

Despite this potential confound, Knösche and colleagues (2005) report that no differences were found between the two baselines, and report primarily on this closer baseline.

Table 3: Music TWs & Baselines



Music			
Effect	**Boundary**	**Music CPS**	**Cadence Effect**
Baselines	-2000 to -1800	-2000 to -1800	-2000 to -1800
	-1800 to -600	-200 to 0	
TWs	-550 to 0	+450 to +600	-1600 to -1300
	-550 to -450		-1200 to -600
	-450 to 0		

In the present study, multiple baseline comparisons were also used. (see Table 3 for a complete table of music TWs and baselines) In order to replicate the procedure of the previous studies, two 200 ms baselines were used. One was taken from the 200 ms preceding the onset of the phrase-initial note following the phrase boundary (i.e. 200 ms during the phrase boundary). Another was taken from the 200 ms preceding the phrase-final note prior to the phrase boundary. This point was chosen because it is the last point where all three conditions were identical. In the stimuli for the present study, this baseline is represented as -2000 to -1800 ms. Further, to investigate the ERP effects during the phrase boundary, a baseline of -1800 to -600 was also chosen. This baseline corresponds with a phrase-final half note. The differences between conditions during this period are a matter of note decay, with the notes in the Phrased and UnPhrased conditions decaying at approximately the same rate, while the held note in the Phrase-NoPause condition decays at a slower rate. Following this note is a rest, in the Phrased condition, a continuation of the note in the NoPause condition, and phrase-boundary filling notes in the UnPhrased condition.

Time windows (Table 3) examined for the music condition included: 450 to 600 ms following the onset of the phrase-initial note after the phrase boundary; the

550 ms preceding the onset of this note (during the phrase boundary); and -1600 to -1300, and -1200 to -600 ms relative to this note based on visual inspection of the ERP waveforms. The 450 to 600 ms time window was used to examine the previously described music CPS. The 550 ms during the phrase boundary were used to examine whether an ERP effect was elicited during the phrase boundary more closely resembling that found in language, i.e. a centro-parietally maximal positivity with sustained duration occurring during the phrase boundary. This time window was further broken down to -550 to -450 and -450 to 0 in order to determine if the scalp distribution of the visually observed effect was due to a P2 like response to the offset of the tone, combined with a positivity of more lasting duration, or a single effect with a consistent distribution.

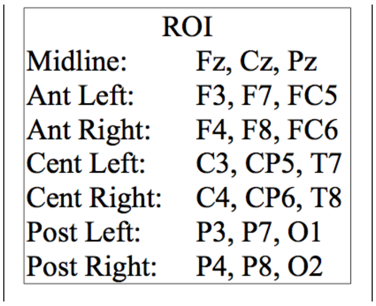
2.5.3 Analysis of Behavioral Data

Statistical analysis of behavioral data was conducted using the statistical analysis software *R* (R Foundation for Statistical Computing, Vienna, Austria), as well as SPSS (Version 22, IBM, Armonk, USA) Behavioral responses were averaged by participant and condition, and then analyzed using repeated measures analysis of variance (ANOVA). Comparisons using 2 x 3 ANOVAs were conducted, with the between-subject factor Grp (ASD x NT), and the three level within-subject factor Cond composed of the three language (A x B x C) and music (Phrased x UnPhrased x Phrased No-Pause) conditions, respectively. Post-hoc comparisons between all condition pairs by domain for each group were then performed using Welch's Two-Sample *t*-tests.

2.5.4 Statistical Analysis of EEG Data

EEG data were broken down into regions of interest and averaged across electrodes for each condition and participant. These regions of interest (ROIs) are summarized in Table 4. The ROIs were composed of 21 electrodes, with 3 in each ROI: Left Anterior F3, F7, FC5; Right Anterior F4, F8, FC6; Left Central C3, CP5, T7; Right Central C4, CP6, T8; Left Posterior P3, P7, O1; Right Posterior P4, P8, O2; and Midline Fz, Cz, Pz.

Table 4: Regions of Interest (ROIs)



ROI	
Midline:	Fz, Cz, Pz
Ant Left:	F3, F7, FC5
Ant Right:	F4, F8, FC6
Cent Left:	C3, CP5, T7
Cent Right:	C4, CP6, T8
Post Left:	P3, P7, O1
Post Right:	P4, P8, O2

EEG data were also split based on music and language conditions. First, responses to each Condition were averaged for each of the TWs and ROIs. Then, repeated measures ANOVAs were conducted for the 600 ms TWs corresponding to the IPh, and each music TW for both lateral and midline ROIs. For the lateral electrodes, a 3 x 2 x 3 x 2 ANOVA containing the factors Cond (3 levels: A x B x C for language; Phrased x NoPause x UnPhrased for music), Hem (2 levels: L x R), AntPost (3 levels: Ant x Cent x Post), as well as the between-subjects factor of Group (2 levels: ASD x NT) was conducted. For the Midline electrodes, a similar repeated measures 3 x 3 x 2 ANOVA was conducted without the factor Hem and using electrode site as the levels for the factor AntPost (Fz x Cz x Pz). Additional ANOVAs were conducted for both Midline and Lateral sites incorporating the factor TW for language CPS investigation in order to determine latency and duration effects of the IPh related positivity. This factor contained 6 levels, one for each of the 100 ms TWs comprising the duration of the phrase boundary. Pairwise comparisons were then conducted between conditions in order to determine which conditions were responsible for significant differences in the global ANOVAs. Greenhouse-Geisser corrections for sphericity were applied where appropriate. Uncorrected degrees of freedom are shown, accompanied by corrected p values.

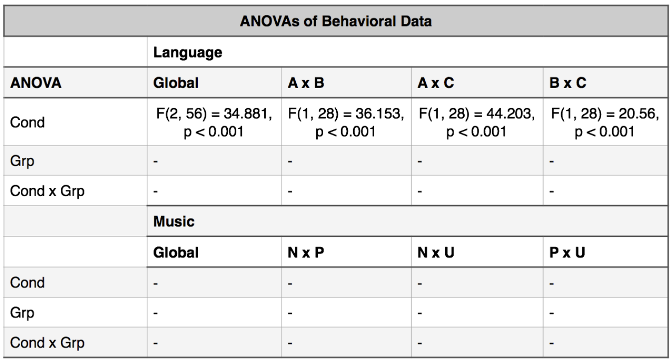
For the Garden Path (N400/P600) effects observed in previous language CPS studies (Steinhauer, et al., 1999; Itzhak, et al., 2010) a peak-to-peak analysis was conducted by filtering the EEG data with a 5 Hz low-pass filter. Then, local minima and maxima were found for the time windows of 250 to 600 ms and 600 to 1200 ms, respectively, at electrodes Cz and Pz following the onset of the crucial second verb of the sentences. The prosody leading up to this verb served as a disambiguating factor for the grammaticality of the sentences. In condition C, the onset of the verb 'to sleep' in **Kevin promises to* ***sleep*** *Anna* served as the trigger for the expected violation. For this time window, using baseline correction was not appropriate given the between condition variability of the stimuli, but by examining the differences between peak amplitudes in the time-windows corresponding to the bi-phasic N400/P600 pattern, it was possible to examine the electro-physical response to determine if the distances between peaks varied significantly between conditions, signifying the presence of these components.

3 Results

3.1 Behavioral Results

The behavioral results for the language and music experiments are summarized in Figures 6 and 7 respectively, with the results of the Analyses of Variance being shown in Table 5. Results of the post-hoc pairwise comparisons across language conditions using Welch's Two Sample t-test are summarized in Table 6, with effect sizes reported using Cohen's d. Only p values below 0.05 are considered significant, and are marked with '*', while p values < 0.01 are marked with '**', and p values < 0.001 are marked with '***'. As can be seen from the Table 5, there was a significant main-effect of Cond for the global ANOVA of all three conditions. There were also highly significant differences in the naturalness ratings of the pairwise comparisons of all three language conditions for all subjects. Condition A was ranked as most natural (Mean: 3.702 (0.66)), while condition B was ranked as less natural than condition A (Mean: 3.194(0.67)). Condition C, which contained the violation, was ranked least natural of the three (Mean: 2.603(0.77)), as predicted. Differences between ratings for each pair of conditions were highly significant, all at the level of $p < 0.001$. No main-effects of Grp were observed, however, nor did any interactions of Cond x Grp reach significance.

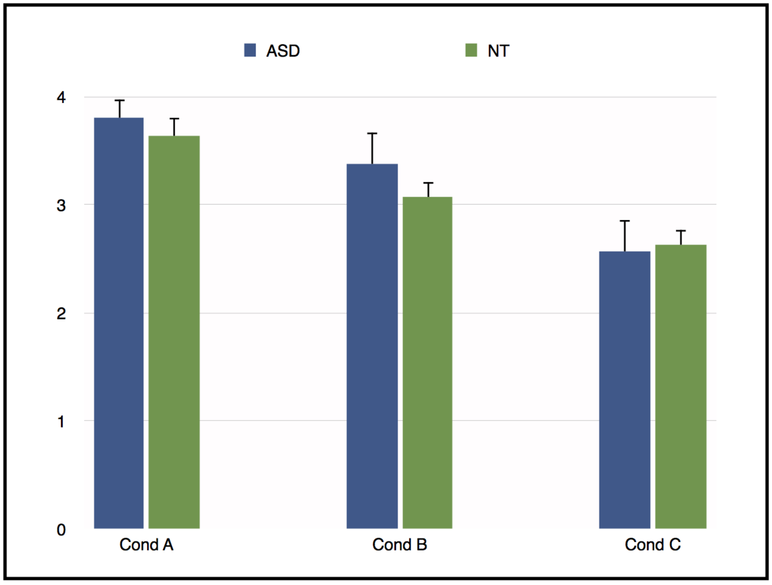
Table 5: Behavioral Results: ANOVAs



ANOVAs of Behavioral Data				
	Language			
ANOVA	**Global**	**A x B**	**A x C**	**B x C**
Cond	F(2, 56) = 34.881, p < 0.001	F(1, 28) = 36.153, p < 0.001	F(1, 28) = 44.203, p < 0.001	F(1, 28) = 20.56, p < 0.001
Grp	-	-	-	-
Cond x Grp	-	-	-	-
	Music			
	Global	**N x P**	**N x U**	**P x U**
Cond	-	-	-	-
Grp	-	-	-	-
Cond x Grp	-	-	-	-

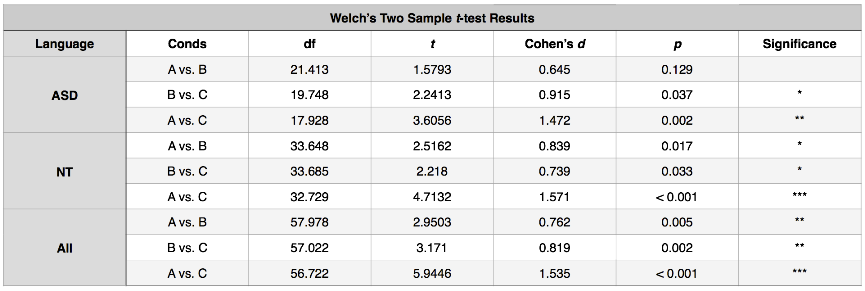
In post-hoc pairwise t-test comparisons of the responses to language conditions,

Figure 6: Behavioral Results: Language



the NT group rated all three groups significantly differently in terms of naturalness, while for the ASD group there was no significant difference between conditions A and B (t $= 1.573$, $p > 0.1$), despite a medium effect size ($d = 0.645$). After combining the groups, however, and repeating the tests, the overall significance of the differences between these two conditions was greater than either group alone (t $= 2.9503$, $p < .01$). This increase in significance for both groups combined and the medium effect size suggest that a significant difference may emerge given a larger ASD group sample. Additionally, between-group comparisons of responses to the behavioral task, using both ANOVA and Welch's *t*-test did not show any significant differences, either by group, or for the Cond x Grp interaction.

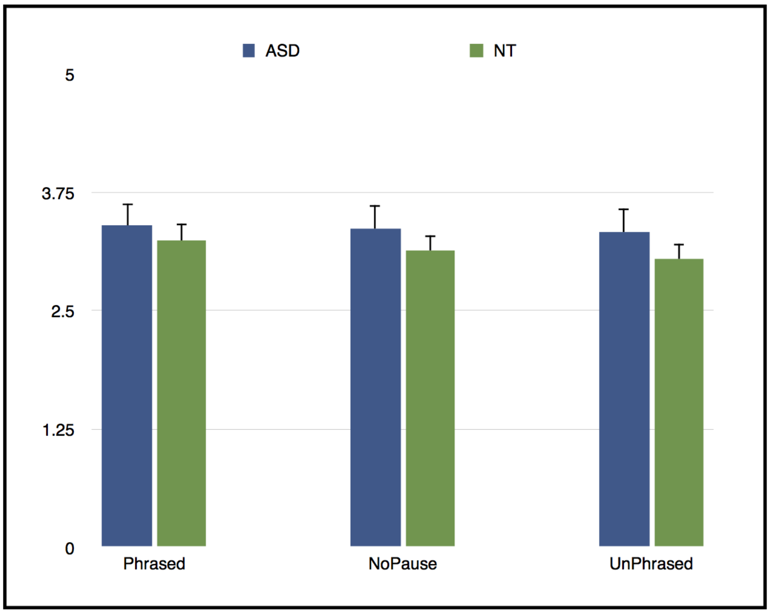
Table 6: Pairwise Comparisons of Language Responses



Welch's Two Sample *t*-test Results						
Language	**Conds**	**df**	***t***	**Cohen's *d***	***p***	**Significance**
ASD	A vs. B	21.413	1.5793	0.645	0.129	
	B vs. C	19.748	2.2413	0.915	0.037	*
	A vs. C	17.928	3.6056	1.472	0.002	**
NT	A vs. B	33.648	2.5162	0.839	0.017	*
	B vs. C	33.685	2.218	0.739	0.033	*
	A vs. C	32.729	4.7132	1.571	< 0.001	***
All	A vs. B	57.978	2.9503	0.762	0.005	**
	B vs. C	57.022	3.171	0.819	0.002	**
	A vs. C	56.722	5.9446	1.535	< 0.001	***

For the ratings of the naturalness of the music stimuli, there were no differences between conditions for either group, nor were there any between group differences. As none of the musical stimuli contained violations, or were somehow unacceptable in terms of musical grammar, the behavioral responses in this domain are unsurprising. In terms of behavioral response, we cannot reject the null hypothesis of there being no between-group differences, for either language or music responses.

Figure 7: Behavioral Results: Music



3.2 EEG Results

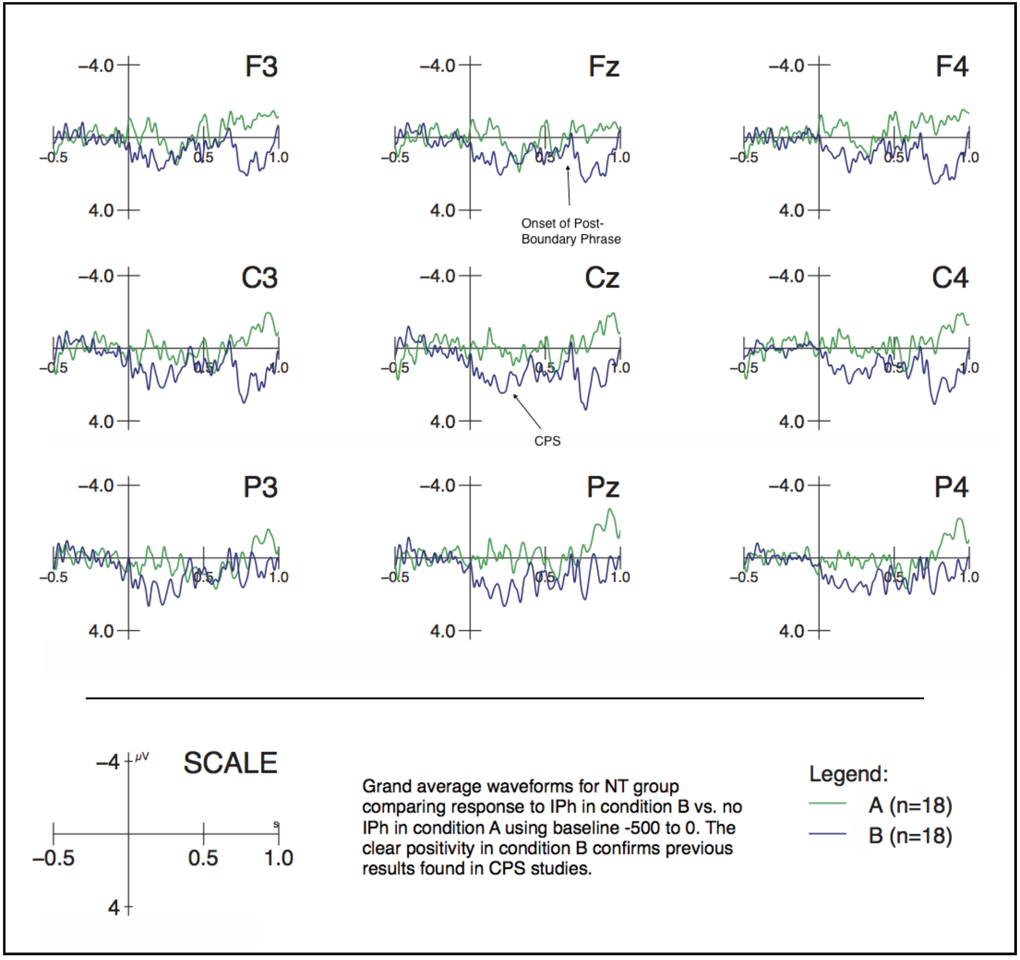
3.2.1 Language

The EEG results of the present experiment are summarized in Figures 8 through 13 and Tables 8 through 12. As with the behavioral results, results with $p < 0.1$ are shown, yet only p values below 0.05 are considered significant, and are marked with '*', while p values < 0.01 are marked with '**', and p values < 0.001 are marked with '***'. Effect sizes are reported for all pairwise comparison main-effects of Cond using Cohen's d.

CPS1: -500 to 0 Baseline

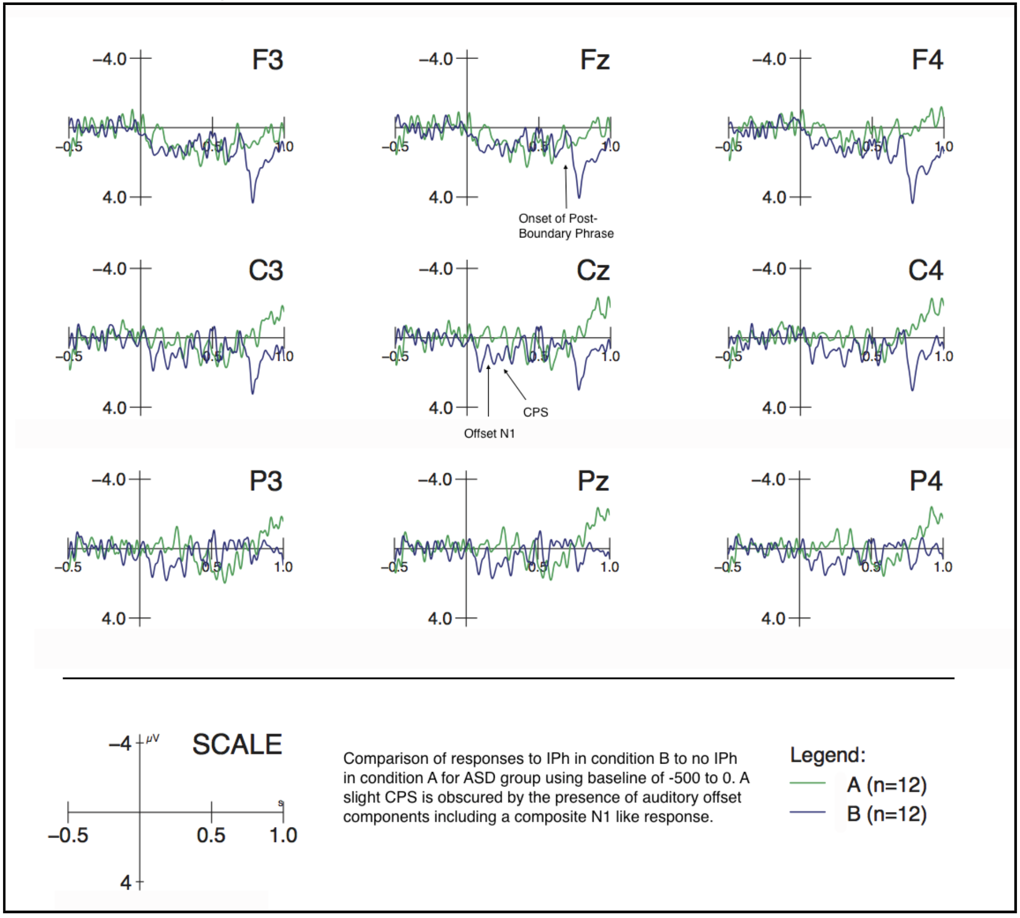
As can be seen from the EEG waveforms in Figure 8, there is a clear positivity (CPS) for the NT group in response to condition B corresponding to the pause at the first phrase boundary relative to condition A where there is no phrase boundary. This figure uses a baseline of -500 to 0 ms, which encompasses the majority of the first phrase in this condition, and much of the beginning of the sentence for condition A. This waveform comparison shows the grand average responses only of the NT group (n = 18). We see here a clear centro-parietal maximality, yet with a wide distribution, confirming the results of previous CPS experiments. The grand average waveforms for the ASD group can be seen in Figure 9.

Figure 8: NT Group CPS1: Conditions A v B



Grand average waveforms for NT group comparing response to IPh in condition B vs. no IPh in condition A using baseline -500 to 0. The clear positivity in condition B confirms previous results found in CPS studies.

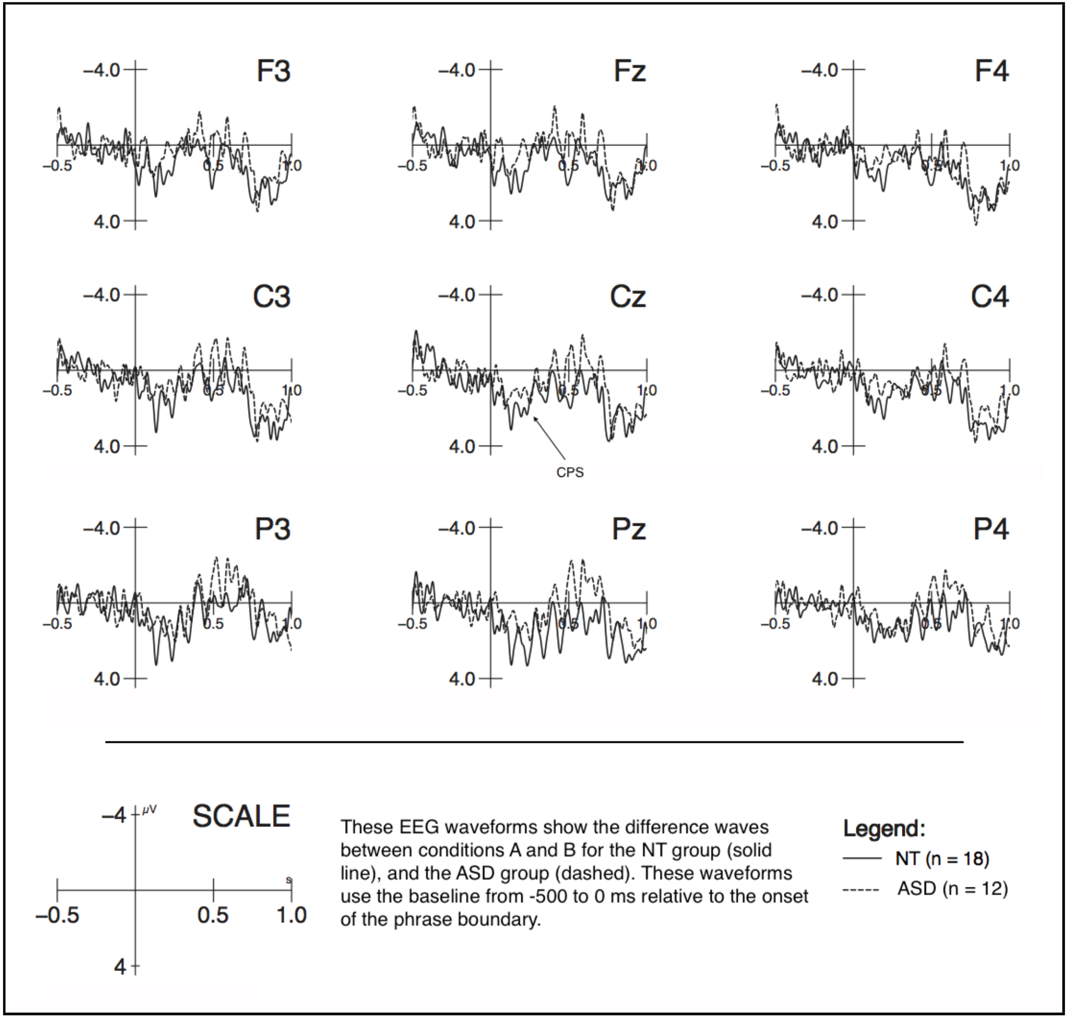
Figure 9: ASD Group CPS1: Conditions A v B



Comparison of responses to IPh in condition B to no IPh in condition A for ASD group using baseline of -500 to 0. A slight CPS is obscured by the presence of auditory offset components including a composite N1 like response.

Comparing the patterns of results by group, the NT group showed a clear CPS, while the ASD group showed a CPS with diminished amplitude that was obscured by auditory offset processing components. The between-group differences between ASD and NT in the CPS TWs can be seen from the difference waves shown in Figure 10. The clear positivity for the NT group contrasts with the less distinct positivity for the ASD group. The positive going shift in the ASD group is of smaller amplitude, and does not have as wide of a distribution as for the NT group. A negative peak occurring around 100-150 ms can also be seen during the CPS TWs corresponding to an offset N1 effect.

Figure 10: ASD vs. NT Difference waves CPS1: Conditions A v B

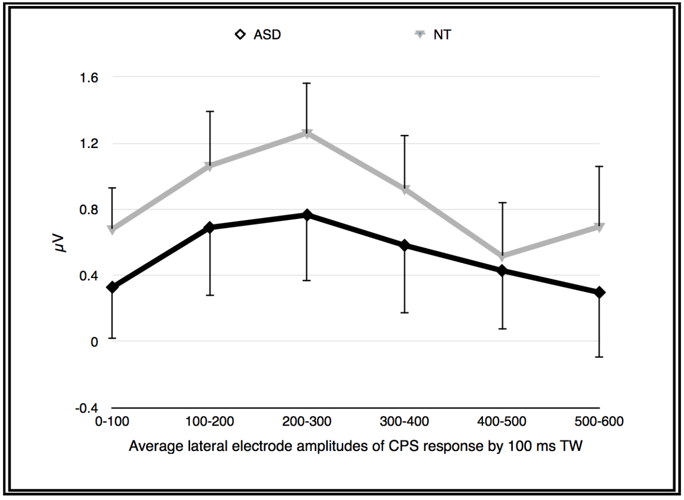


These EEG waveforms show the difference waves between conditions A and B for the NT group (solid line), and the ASD group (dashed). These waveforms use the baseline from -500 to 0 ms relative to the onset of the phrase boundary.

Legend:
—— NT (n = 18)
----- ASD (n = 12)

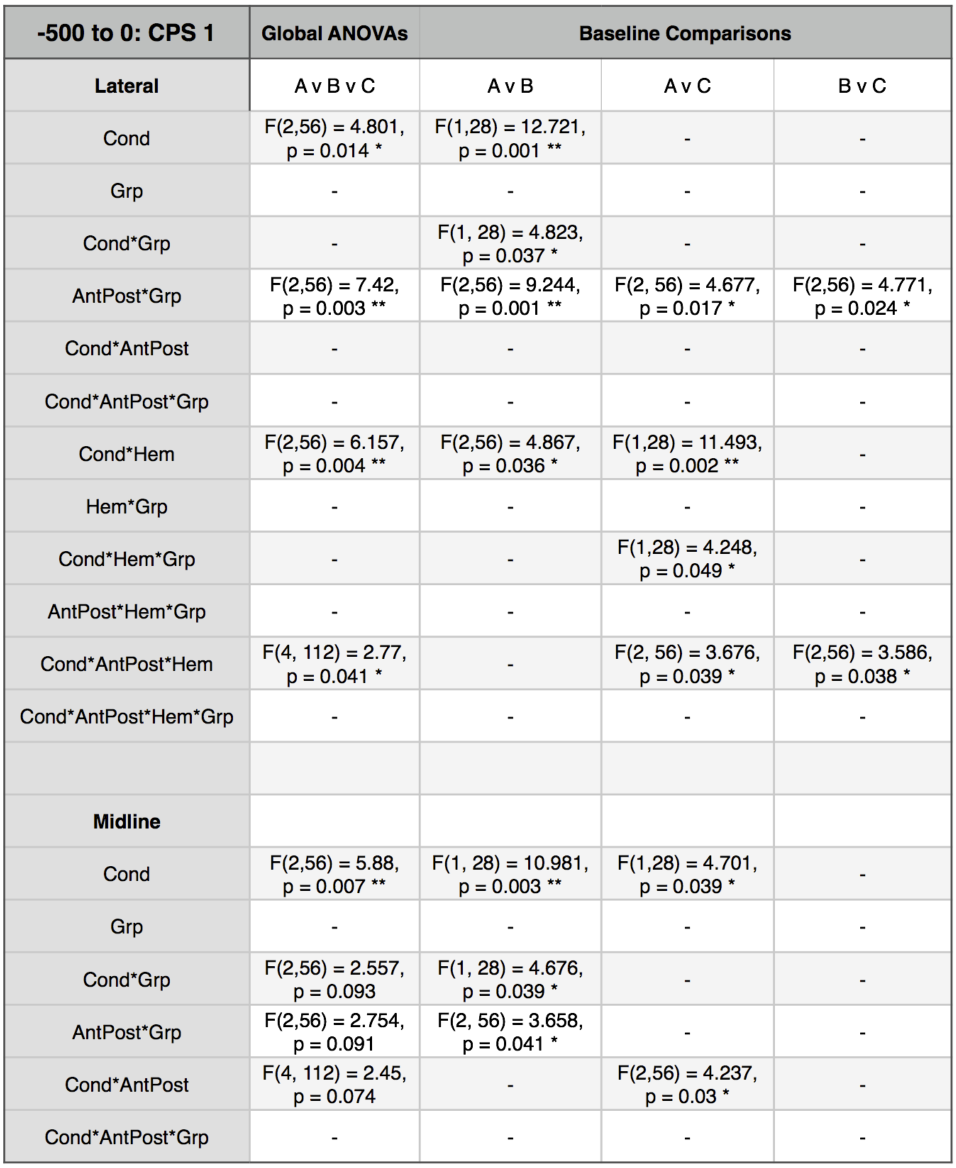
These results are further confirmed by 100 ms TW averages of Cond B response across all lateral electrodes as shown in Figure 11. The diminished positivity for the ASD group is clear. Together, these results suggest that the prosody dependent chunking mechanism responsible for consolidating the previous IPh is less robust for this group.

Figure 11: Condition B Responses by 100 ms TW



For the ANOVAs conducted for the first IPh, please refer to Table 7. For the Global ANOVAs in the CPS1 TW, a significant main-effect of the factor Cond was observed, both at lateral sites ($F(2{,}56) = 4.801$, $p = 0.014$), and at midline sites ($F(2{,}56) = 5.88$, $p = 0.007$). These effects show that electrophysiological responses to the presence vs. absence of a phrase boundary were significantly different when all subjects were taken into account. Additionally, a trend toward a significant interaction of Cond x Grp was observed at midline sites ($F(2{,}56) = 2.557$, $p = 0.093$), providing evidence for a between group difference in response to prosodic phrase boundaries. This effect was not significant at lateral sites, however. There was also a significant interaction of AntPost x Grp at lateral sites ($p = 0.003$), with a trend toward significance for midline sites ($p = 0.091$). This was due to a more anterior distribution of effects in the ASD group across conditions as compared to the NT group whose responses were more posterior. A comparison of the anterior/posterior distribution of responses to condition B (CPS) can be seen in Figure 12. An interaction of the factors Cond x Hem ($F(2,\ 56) = 4.898$, $p = 0.035$) at lateral sites was also observed due to a greater response over the right hemisphere for both groups (see -50 to +50 baseline section for details).

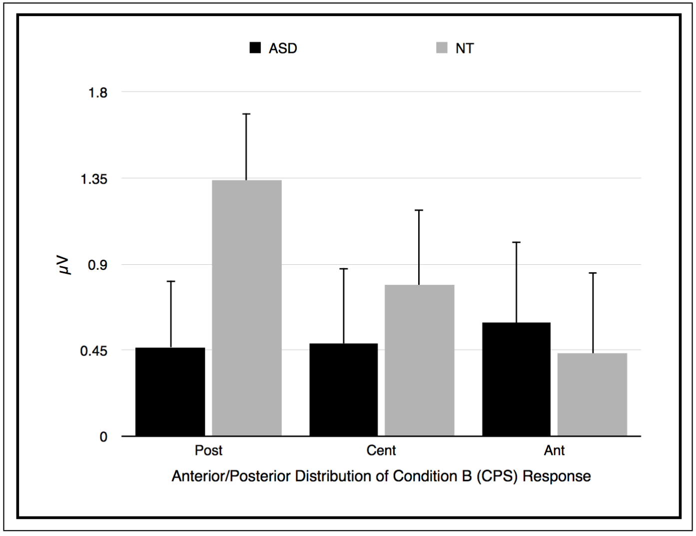
Table 7: ANOVAs: -500 to 0 Baseline CPS1



-500 to 0: CPS 1	**Global ANOVAs**	**Baseline Comparisons**		
Lateral	A v B v C	A v B	A v C	B v C
Cond	F(2,56) = 4.801, p = 0.014 *	F(1,28) = 12.721, p = 0.001 **	-	-
Grp	-	-	-	-
Cond*Grp	-	F(1, 28) = 4.823, p = 0.037 *	-	-
AntPost*Grp	F(2,56) = 7.42, p = 0.003 **	F(2,56) = 9.244, p = 0.001 **	F(2, 56) = 4.677, p = 0.017 *	F(2,56) = 4.771, p = 0.024 *
Cond*AntPost	-	-	-	-
Cond*AntPost*Grp	-	-	-	-
Cond*Hem	F(2,56) = 6.157, p = 0.004 **	F(2,56) = 4.867, p = 0.036 *	F(1,28) = 11.493, p = 0.002 **	-
Hem*Grp	-	-	-	-
Cond*Hem*Grp	-	-	F(1,28) = 4.248, p = 0.049 *	-
AntPost*Hem*Grp	-	-	-	-
Cond*AntPost*Hem	F(4, 112) = 2.77, p = 0.041 *	-	F(2, 56) = 3.676, p = 0.039 *	F(2,56) = 3.586, p = 0.038 *
Cond*AntPost*Hem*Grp	-	-	-	-
Midline				
Cond	F(2,56) = 5.88, p = 0.007 **	F(1, 28) = 10.981, p = 0.003 **	F(1,28) = 4.701, p = 0.039 *	-
Grp	-	-	-	-
Cond*Grp	F(2,56) = 2.557, p = 0.093	F(1, 28) = 4.676, p = 0.039 *	-	-
AntPost*Grp	F(2,56) = 2.754, p = 0.091	F(2, 56) = 3.658, p = 0.041 *	-	-
Cond*AntPost	F(4, 112) = 2.45, p = 0.074	-	F(2,56) = 4.237, p = 0.03 *	-
Cond*AntPost*Grp	-	-	-	-

Pairwise comparisons of this larger TW show that when comparing just conditions A x B, the conditions without and with the phrase boundary, respectively, the effect of Cond is more significant than that seen in the results of the global ANOVA at both lateral ($F(1,28) = 12.721$, $p = 0.001$, $d = 0.621$) and midline ($F(1, 28) = 10.981$, $p = 0.003$, $d = 0.757$) sites. Additionally, the Cond x Grp interaction reached significance for both ROI comparisons, both at the level of $p < 0.05$. This comparison provides further support for the hypothesis of between-

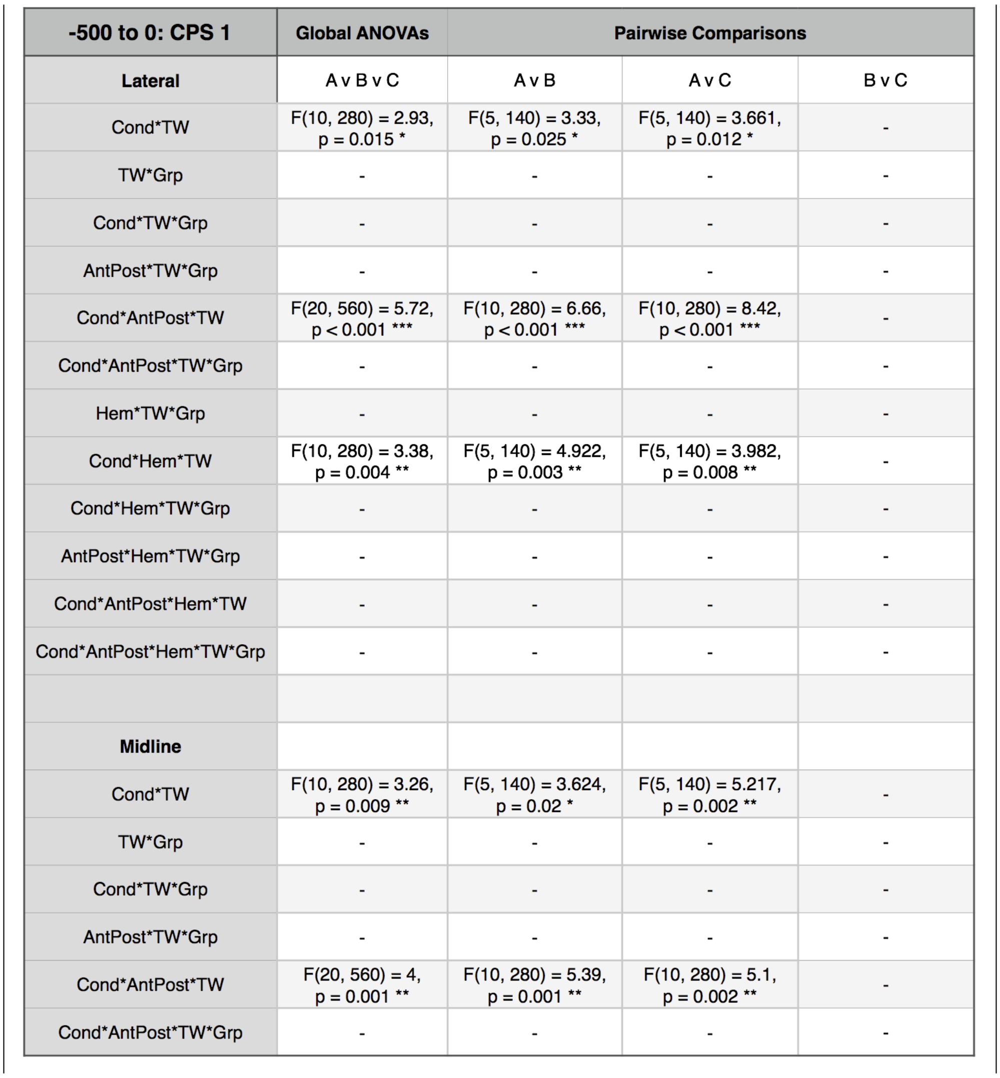
Figure 12: Anterior/Posterior Distribution of Condition B Responses



group differences in prosodic boundary processing. No significant effects of Cond or interactions of Cond x Grp were observed for the comparison of conditions A x C, however, which is somewhat surprising given the fact that conditions B and C are identical during this period.

Interactions with the factor TW were also informative (see Table 8). A significant interaction of Cond x TW was observed at both lateral sites ($F(10,\ 280) = 3.33$, $p = 0.025$) and midline sites ($F(10,\ 280) = 3.26$, $p = 0.009$), which simply shows that the effect of Condition changes over time. Additionally, highly significant interactions of Cond x AntPost x TW were observed at both lateral ($F(20,\ 560) = 5.72$, $p < 0.001$) and midline ($F(10,\ 280) = 4.0$, $p = 0.001$) sites, suggesting that in the earlier stages of the CPS effect, it is more broadly distributed due to its size, while as it begins to decay, the central/posterior distribution becomes more apparent. This was confirmed by an interaction of Cond x Hem x TW ($F(10,\ 280) = 3.38$, $p = 0.004$), due to the positivity being prolonged over the right hemisphere, which is largely considered responsible for the prosodic processing pathway. No interaction was observed between the factors of TW x Grp, meaning that while it appears from the ERP waveforms that the ASD group's response has a shorter duration, this interaction was not significant.

Table 8: ANOVAs: -500 to 0 Baseline CPS1 x TW



-500 to 0: CPS 1	**Global ANOVAs**	**Pairwise Comparisons**		
Lateral	A v B v C	A v B	A v C	B v C
Cond*TW	F(10, 280) = 2.93, p = 0.015 *	F(5, 140) = 3.33, p = 0.025 *	F(5, 140) = 3.661, p = 0.012 *	-
TW*Grp	-	-	-	-
Cond*TW*Grp	-	-	-	-
AntPost*TW*Grp	-	-	-	-
Cond*AntPost*TW	F(20, 560) = 5.72, p < 0.001 ***	F(10, 280) = 6.66, p < 0.001 ***	F(10, 280) = 8.42, p < 0.001 ***	-
Cond*AntPost*TW*Grp	-	-	-	-
Hem*TW*Grp	-	-	-	-
Cond*Hem*TW	F(10, 280) = 3.38, p = 0.004 **	F(5, 140) = 4.922, p = 0.003 **	F(5, 140) = 3.982, p = 0.008 **	-
Cond*Hem*TW*Grp	-	-	-	-
AntPost*Hem*TW*Grp	-	-	-	-
Cond*AntPost*Hem*TW	-	-	-	-
Cond*AntPost*Hem*TW*Grp	-	-	-	-
Midline				
Cond*TW	F(10, 280) = 3.26, p = 0.009 **	F(5, 140) = 3.624, p = 0.02 *	F(5, 140) = 5.217, p = 0.002 **	-
TW*Grp	-	-	-	-
Cond*TW*Grp	-	-	-	-
AntPost*TW*Grp	-	-	-	-
Cond*AntPost*TW	F(20, 560) = 4, p = 0.001 **	F(10, 280) = 5.39, p = 0.001 **	F(10, 280) = 5.1, p = 0.002 **	-
Cond*AntPost*TW*Grp	-	-	-	-

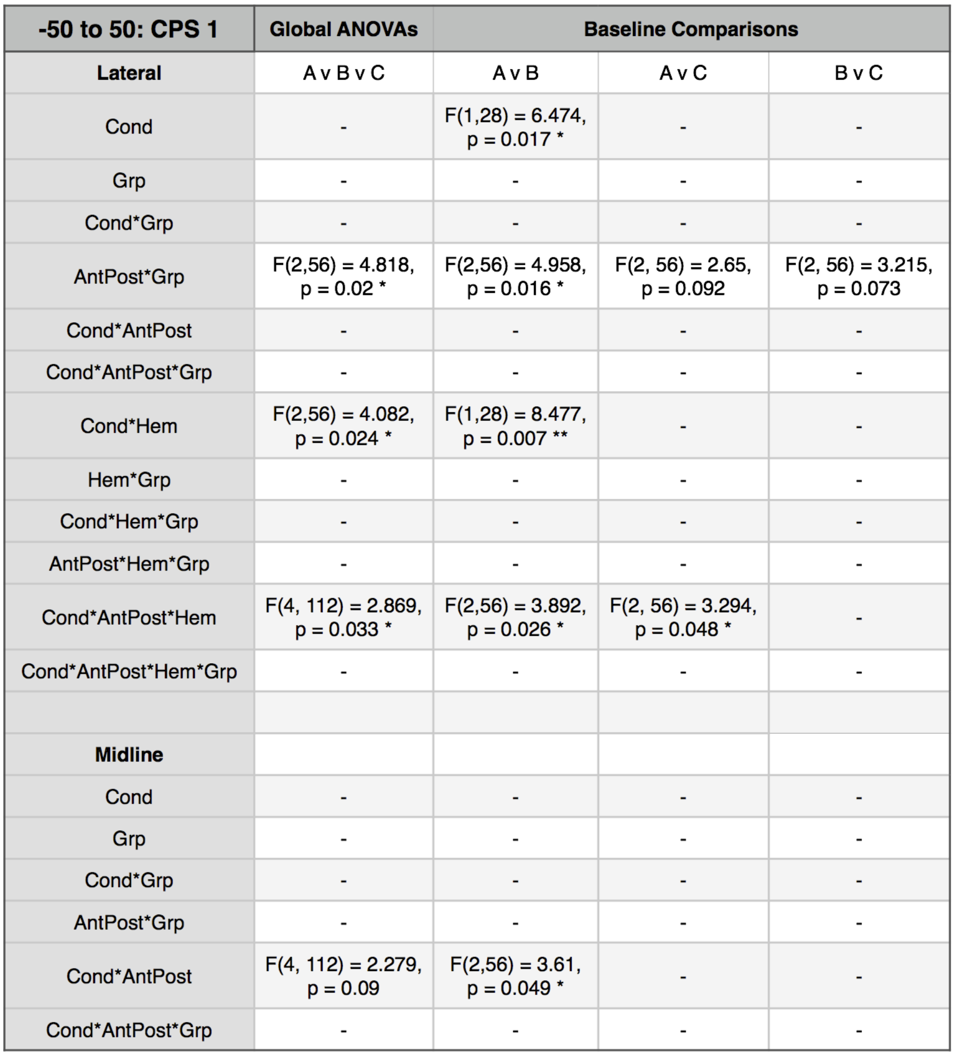
Pairwise comparisons of conditions A x B and A x C show similar TW interactions as those seen in the global ANOVA for lateral and midline sites. No pairwise comparisons between B x C resulted in significance, however, with the exception of an AntPost x Grp interaction ($F(2, 56) = 6.52$, $p = 0.008$), which was constant across all comparisons, and a Cond x AntPost x Hem interaction. Conditions B and C are acoustically identical during this period, thus, no main-effects of Cond were expected.

CPS1: -50 to +50 Baseline

Results of ANOVAs conducted using the baseline of -50 ms to +50 ms relative to the onset of the pause are shown in Table 9. This baseline was used to confirm the results observed with the longer baseline prior to the onset of the IPh. The fact that it includes 50 ms during which the CPS is observable means that this baseline eliminates some of the robustness of the observed effect, however. This can be seen in the pattern of results which are generally less significant than those corrected using the longer, non-overlapping baseline. No main-effect of Cond was observed for lateral electrodes, nor for midline electrodes in the global ANOVA. Also, importantly, no Cond x Grp interaction reached significance, and while the AntPost x Grp interaction was quite robust for the previous baseline (Global: $p = 0.001$), the interaction decreased in significance here ($F(2, 56) = 4.818$, $p = 0.02$).

For the pairwise comparison of A x B, there was a significant main-effect of Cond ($F(1,28) = 6.473$, $p = 0.017$, $d = 0.538$) at lateral sites, if not at the midline electrodes ($p > 0.1$, $d = 0.256$). There was also a significant Cond x Hem interaction ($F = 8.477$, $p = 0.007$), which helped to drive a Cond x AntPost x Hem interaction ($F = 3.892$, $p = 0.026$). Despite this latter 3 factor interaction, the Cond x AntPost interaction was not significant at lateral sites, however, at midline sites ($F(2,56) = 3.61$, $p = 0.049$) it did reach significance. Taken together, these findings confirm that the phrase-boundary induced positivity has a longer

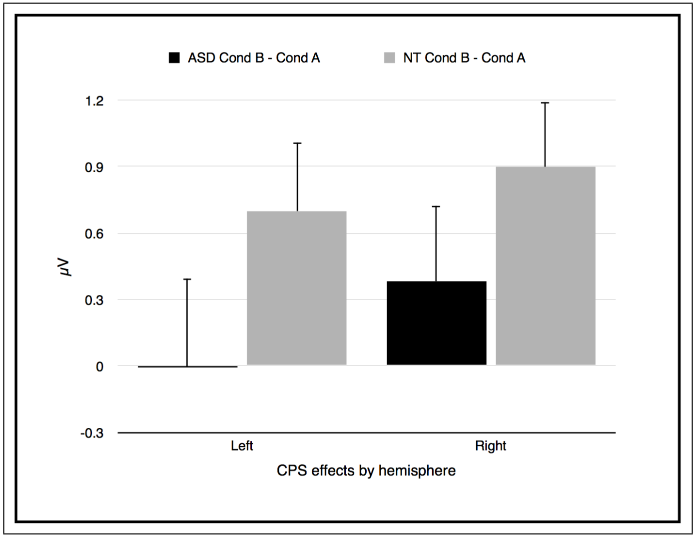
Table 9: ANOVAs: -50 to 50 Baseline CPS1



-50 to 50: CPS 1	Global ANOVAs	Baseline Comparisons		
Lateral	A v B v C	A v B	A v C	B v C
Cond	-	F(1,28) = 6.474, p = 0.017 *	-	-
Grp	-	-	-	-
Cond*Grp	-	-	-	-
AntPost*Grp	F(2,56) = 4.818, p = 0.02 *	F(2,56) = 4.958, p = 0.016 *	F(2, 56) = 2.65, p = 0.092	F(2, 56) = 3.215, p = 0.073
Cond*AntPost	-	-	-	-
Cond*AntPost*Grp	-	-	-	-
Cond*Hem	F(2,56) = 4.082, p = 0.024 *	F(1,28) = 8.477, p = 0.007 **	-	-
Hem*Grp	-	-	-	-
Cond*Hem*Grp	-	-	-	-
AntPost*Hem*Grp	-	-	-	-
Cond*AntPost*Hem	F(4, 112) = 2.869, p = 0.033 *	F(2,56) = 3.892, p = 0.026 *	F(2, 56) = 3.294, p = 0.048 *	-
Cond*AntPost*Hem*Grp	-	-	-	-
Midline				
Cond	-	-	-	-
Grp	-	-	-	-
Cond*Grp	-	-	-	-
AntPost*Grp	-	-	-	-
Cond*AntPost	F(4, 112) = 2.279, p = 0.09	F(2,56) = 3.61, p = 0.049 *	-	-
Cond*AntPost*Grp	-	-	-	-

duration and larger amplitude more posteriorly (at least for the NT group) and over the right hemisphere. This right hemispheric lateralization for both groups confirms the predictions of the dual-pathway processing model where prosodic information is largely processed in the right hemisphere. The effect of hemisphere was especially pronounced for the ASD group, who showed almost no effect of Cond across averaged left-hemisphere sites. The relative hemispheric distribution of effects can be seen in Figure 13, which shows the difference in responses (B – A) to the two conditions by hemisphere and group.

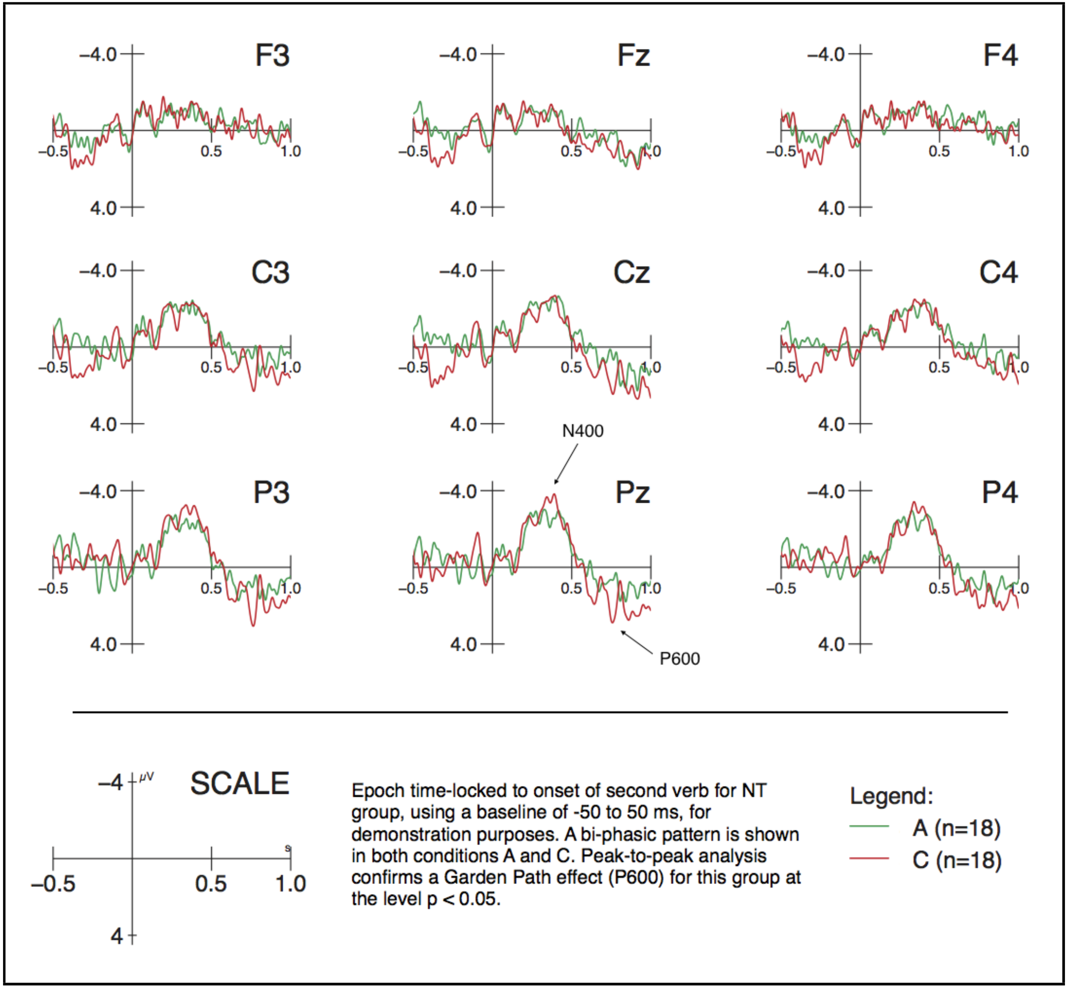
Figure 13: Hemisphere Effects of CPS: Cond B - Cond A



Interactions involving the factor TW tended to confirm previous observations, yet also gave further depth to the comparisons using this baseline, especially as in the first TWs of the phrase boundary the different responses were compressed by the baseline. There were interactions of Cond x TW at lateral ($F(10, 280) = 2.93$, $p = 0.015$), and midline ($F(10, 280) = 3.26$, $p = 0.009$) sites, driven by differences between condition A and B/C. While there was no main-effect of Cond, these interactions, as well as a highly significant Cond x AntPost x TW interaction ($F(10, 280) = 4$, $p = 0.001$) show that over the course of these TWs, the conditions did elicit significantly different responses, and these differed in both Anterior/Posterior and Hemispheric distribution.

Garden Path Effects

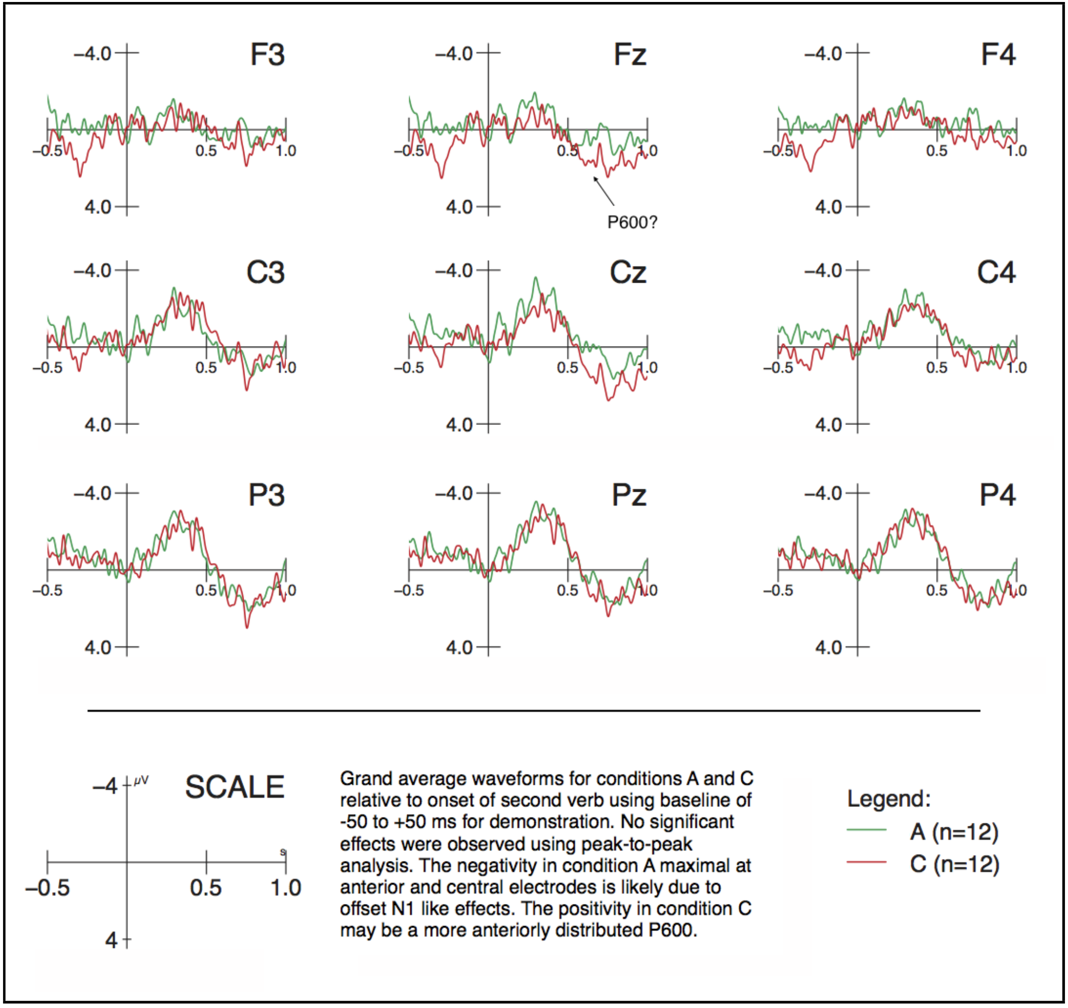
Figure 14: NT Garden Path Response: Conditions A v C



Epoch time-locked to onset of second verb for NT group, using a baseline of -50 to 50 ms, for demonstration purposes. A bi-phasic pattern is shown in both conditions A and C. Peak-to-peak analysis confirms a Garden Path effect (P600) for this group at the level $p < 0.05$.

In the first study to identify the CPS, a spliced condition was also used that contained a violation of argument structure resulting from incongruous prosody. (Steinhauer, et al., 1999) As a response to this violation, bi-phasic N400/P600 like garden path effects were observed. The present study sought to replicate these effects, and used similar prosodic violations. Figure 14 shows the grand average waveforms for the NT group with a comparison between conditions A and C, which are acoustically identical during this period, but condition C contains a violation due to incongruous prosody as a result of splicing. A baseline of -50 to +50 ms relative to the onset of the verb is used for demonstration purposes, as raw peak-to-peak analysis was conducted in the absence of baseline correction. Peak-to-peak analysis finds minima and maxima per participant and TW, and measures the distance between the two.

Figure 15: ASD Group Garden Path Response: Conditions A v C



Grand average waveforms for conditions A and C relative to onset of second verb using baseline of -50 to +50 ms for demonstration. No significant effects were observed using peak-to-peak analysis. The negativity in condition A maximal at anterior and central electrodes is likely due to offset N1 like effects. The positivity in condition C may be a more anteriorly distributed P600.

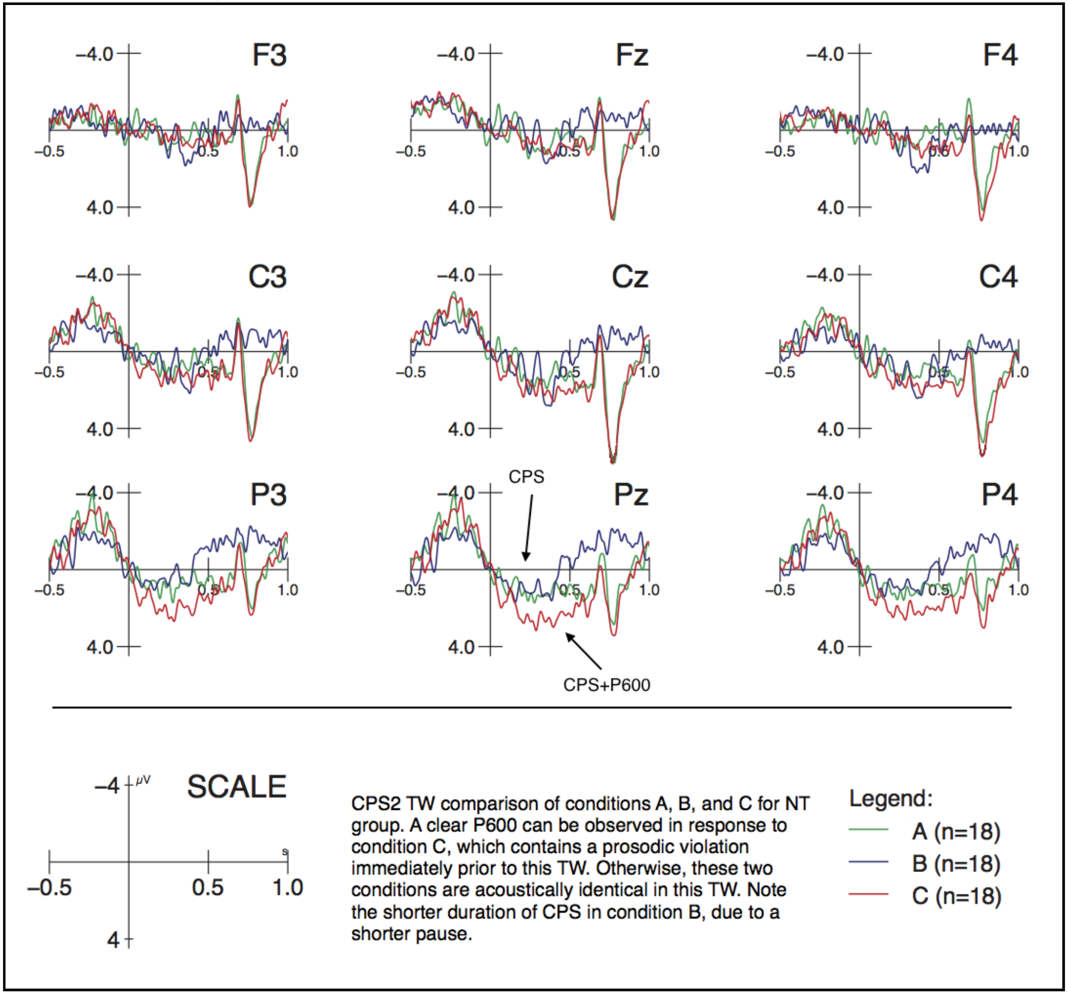
Using this peak-to-peak distance comparison across conditions, a significant main-effect of Cond was observed between conditions A x C for the NT group at electrode Pz ($F(1, 34) = 5.4328$, $p = 0.026$, $d = 0.777$). (see Figure 14) No other significant differences were observed for other ANOVAs conducted, including the NT group at electrode Cz ($p > 0.1$, $d = 0.054$), the ASD group at electrode Pz ($p > 0.1$, $d = 0.003$), and the Cond x Grp interaction at Pz ($p > 0.1$). A slight trend toward significance was also observed for the main-effect of Cond for all participants together at electrode Pz ($F(1, 58) = 2.799$, $p = 0.099$, $d = 0.432$), likely driven by the effect in the NT group. Figure 15 shows the EEG waveforms for the ASD group for conditions A and C using a baseline of -50 to +50 ms relative to the onset of Verb 2 for demonstration purposes.

A clear bi-phasic pattern was observed for all conditions, showing that this bi-phasic pattern is not solely a result of prosodic violation, as the prosodic incongruity only occurs in condition C, but is due to the summation of multiple elements of sentence processing. The positivity observed in the waveforms is at least partially the result of the CPS 2, as the violation occurs immediately prior to the second IPh in condition C, yet the difference between conditions seen in the waveforms and statistical analysis suggests that a P600 is also present. The possible components comprising this bi-phasic pattern will be considered in the Discussion section.

CPS 2

As previously mentioned, the second IPh in conditions B and C corresponded with the first IPh in condition A. Only one baseline was used to analyze the TWs for CPS 2, since all three conditions were different: two (A and C) had the same immediately preceding word and pause length, but one (C) was spliced in the closure during the affricate /ts/ in the onset of the penultimate (relative to the IPh) word 'zu', creating one condition that contained a violation, and one that was naturally spoken. Prior to this point, the spliced condition was identical to condition B, which contains a different, transitive second verb. Thus, all three conditions were different in the 500 ms preceding the onset of the IPh, resulting in the need to use a single proximate baseline. For consistency, and following previous studies, a baseline of -50 to +50 ms relative to the onset of the second pause was chosen.

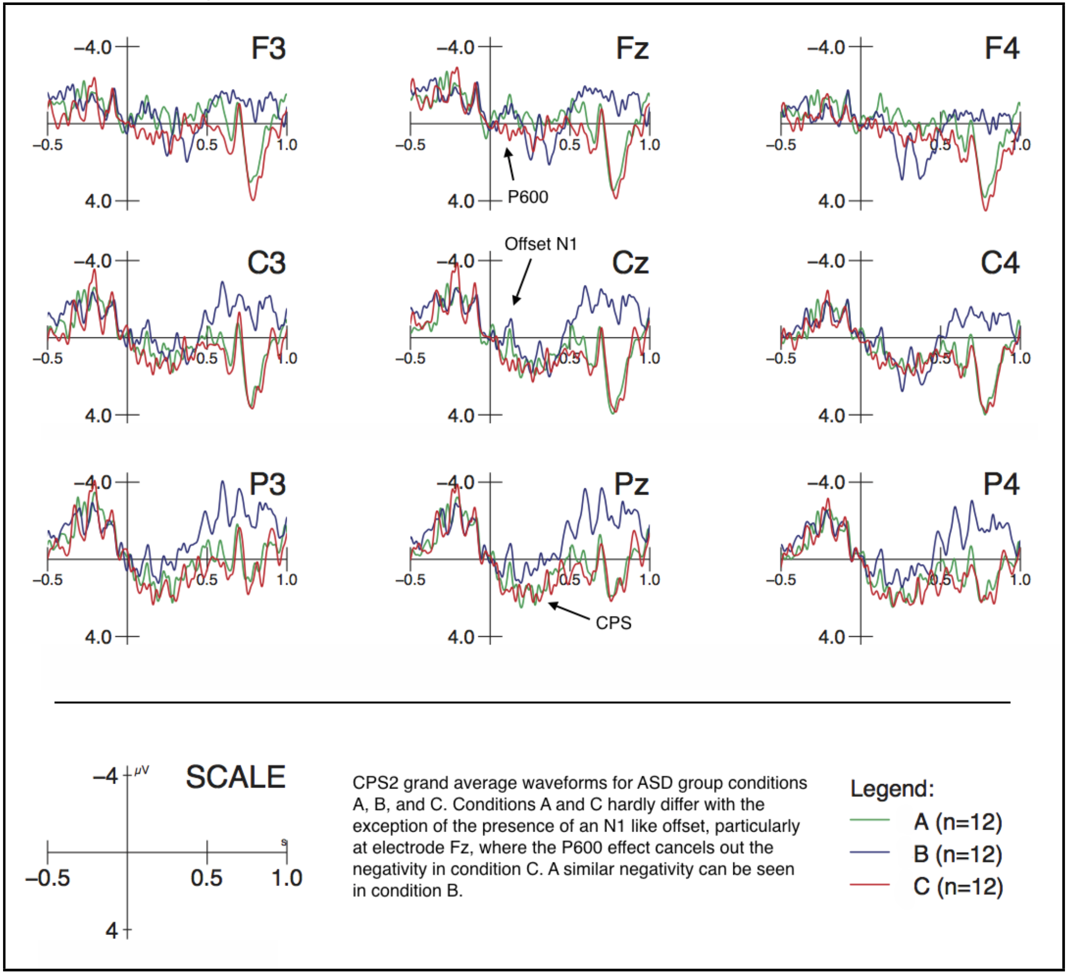
Figure 16: NT Group CPS2: Conditions A v B v C



CPS2 TW comparison of conditions A, B, and C for NT group. A clear P600 can be observed in response to condition C, which contains a prosodic violation immediately prior to this TW. Otherwise, these two conditions are acoustically identical in this TW. Note the shorter duration of CPS in condition B, due to a shorter pause.

The grand averages of the NT group for all three conditions are shown in Figure 16. Since an IPh exists in all 3 conditions, a positive shift can be observed in all three. Note that conditions A and C are acoustically identical during this period, and thus they produce nearly identical waveforms, with the exception that the large positivity at the posterior electrodes in condition C is the result of a summated effect of the CPS 2 and P600. The grand averages of the ASD group for all three conditions are shown in Figure 17. These all show a slight positivity during the prosodic break. An additional N1 offset component can be seen, particularly at more anterior electrodes in condition A and B, which may be canceled out by a P600 in condition C.

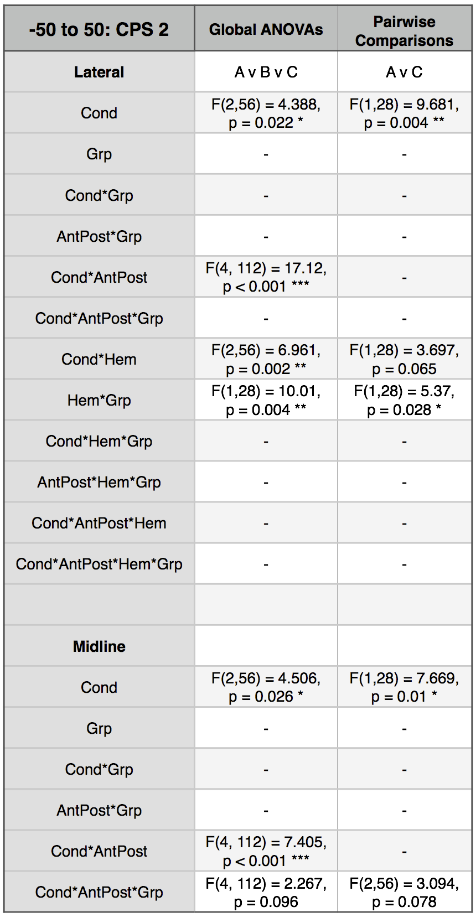
Figure 17: ASD Group CPS2: Conditions A v B v C



CPS2 grand average waveforms for ASD group conditions A, B, and C. Conditions A and C hardly differ with the exception of the presence of an N1 like offset, particularly at electrode Fz, where the P600 effect cancels out the negativity in condition C. A similar negativity can be seen in condition B.

As can be seen from Table 10, the pattern of results differed from that of CPS1. As the phrase boundary was shorter in condition B, differences between responses to this condition and the other two were due to differences in acoustical features of the pause. Therefore, the pairwise comparison between conditions A and C is the only one shown in the statistical table, and, aside from the results of the global ANOVA, will be the only comparison discussed.

Table 10: ANOVAs: -50 to 50 Baseline CPS2



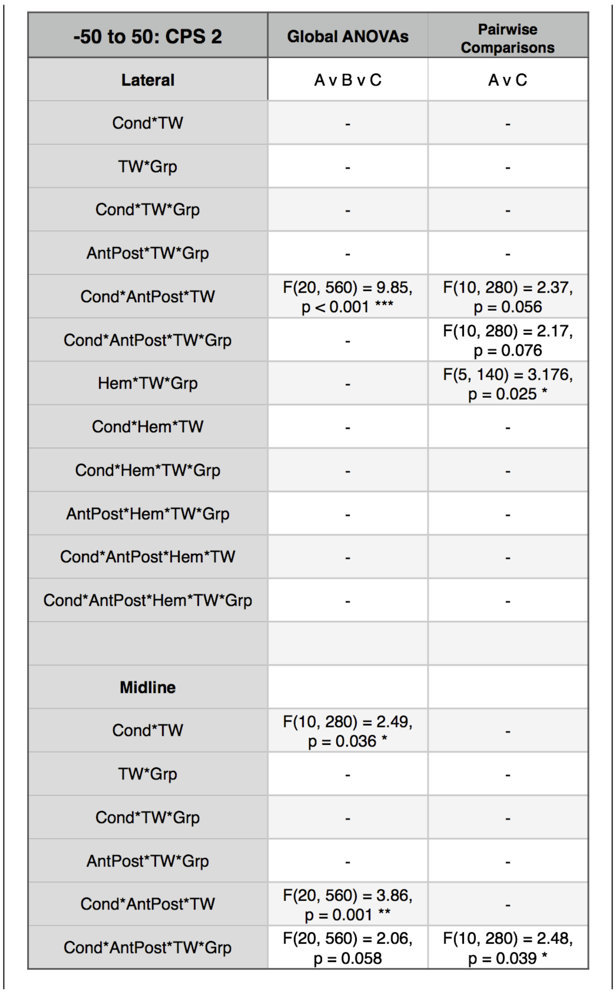
-50 to 50: CPS 2	Global ANOVAs	Pairwise Comparisons
Lateral	A v B v C	A v C
Cond	F(2,56) = 4.388, p = 0.022 *	F(1,28) = 9.681, p = 0.004 **
Grp	-	-
Cond*Grp	-	-
AntPost*Grp	-	-
Cond*AntPost	F(4, 112) = 17.12, p < 0.001 ***	-
Cond*AntPost*Grp	-	-
Cond*Hem	F(2,56) = 6.961, p = 0.002 **	F(1,28) = 3.697, p = 0.065
Hem*Grp	F(1,28) = 10.01, p = 0.004 **	F(1,28) = 5.37, p = 0.028 *
Cond*Hem*Grp	-	-
AntPost*Hem*Grp	-	-
Cond*AntPost*Hem	-	-
Cond*AntPost*Hem*Grp	-	-
Midline		
Cond	F(2,56) = 4.506, p = 0.026 *	F(1,28) = 7.669, p = 0.01 *
Grp	-	-
Cond*Grp	-	-
AntPost*Grp	-	-
Cond*AntPost	F(4, 112) = 7.405, p < 0.001 ***	-
Cond*AntPost*Grp	F(4, 112) = 2.267, p = 0.096	F(2,56) = 3.094, p = 0.078

For the global ANOVA at the lateral electrodes, there was a significant main-effect of the factor Cond both at lateral ($F(2{,}56) = 4.388$, $p = 0.022$) and midline ($F(2{,}56) = 4.506$, $p = 0.026$) electrodes, yet there were not any significant effects of Grp or interactions of Cond x Grp. This effect of Cond in the global ANOVA was driven by the different responses observed between conditions A and C ($F(1{,}28)$

$= 9.681$, $p = 0.004$, $d = 0.554$, at lateral sites), which differ only in the presence of a violation in condition C. Thus, these different responses by condition can be attributed to the summation of CPS (found in both conditions) and P600 (found only in condition C).

A trend toward significance was also observed for the interaction of Cond x Hem for the pairwise A x C ANOVA ($F = 3.697$, $p = 0.065$), as was a significant Hem x Grp interaction, in both global ($F(1,28) = 10.01$, $p = 0.004$) and pairwise comparisons ($F(1,28) = 5.37$, $p = 0.028$). These interactions were due to the ASD group having an almost entirely right-lateralized CPS response in both conditions, while the additional positivity in response to condition C was more left-lateralized.

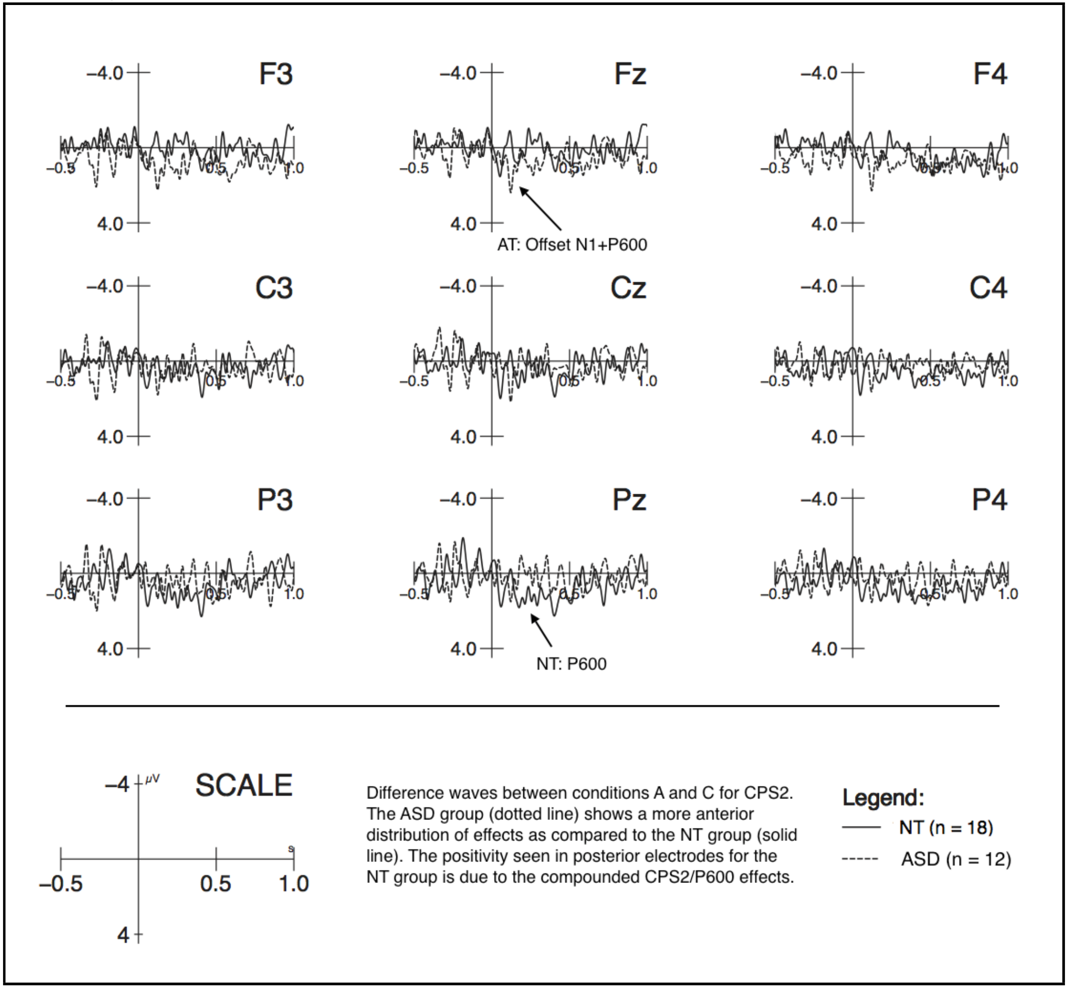
Table 11: ANOVAs: -50 to 50 Baseline CPS2 x TW



-50 to 50: CPS 2	**Global ANOVAs**	**Pairwise Comparisons**
Lateral	A v B v C	A v C
Cond*TW	-	-
TW*Grp	-	-
Cond*TW*Grp	-	-
AntPost*TW*Grp	-	-
Cond*AntPost*TW	F(20, 560) = 9.85, p < 0.001 ***	F(10, 280) = 2.37, p = 0.056
Cond*AntPost*TW*Grp	-	F(10, 280) = 2.17, p = 0.076
Hem*TW*Grp	-	F(5, 140) = 3.176, p = 0.025 *
Cond*Hem*TW	-	-
Cond*Hem*TW*Grp	-	-
AntPost*Hem*TW*Grp	-	-
Cond*AntPost*Hem*TW	-	-
Cond*AntPost*Hem*TW*Grp	-	-
Midline		
Cond*TW	F(10, 280) = 2.49, p = 0.036 *	-
TW*Grp	-	-
Cond*TW*Grp	-	-
AntPost*TW*Grp	-	-
Cond*AntPost*TW	F(20, 560) = 3.86, p = 0.001 **	-
Cond*AntPost*TW*Grp	F(20, 560) = 2.06, p = 0.058	F(10, 280) = 2.48, p = 0.039 *

Despite interactions in the global ANOVA (Table 11) between Cond x AntPost, Cond x TW, and Cond x AntPost x TW, these differences were due to interactions involving condition B, and thus, are due to differences in acoustic features of stimuli. A trend toward a significant interaction of the factors Cond x AntPost x TW x Grp in the global ANOVA, however, was driven by a significant interaction of these factors between conditions A and C ($F(10, 280) = 2.481$, $p = 0.039$), likely due to an offset N1 effect observed only in the ASD group, which is cancelled out in condition C by a more anterior distribution of the P600 in this group.

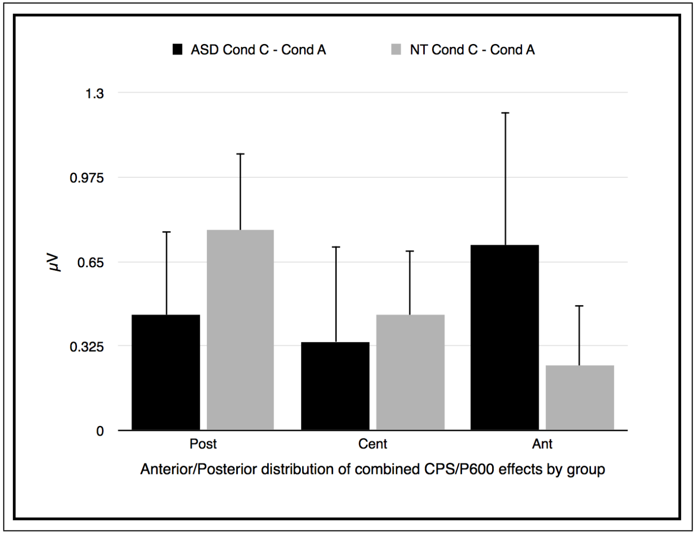
Figure 18: ASD vs. NT Difference Waves CPS2: Conditions A v C



Difference waves between conditions A and C for CPS2. The ASD group (dotted line) shows a more anterior distribution of effects as compared to the NT group (solid line). The positivity seen in posterior electrodes for the NT group is due to the compounded CPS2/P600 effects.

Examining the difference waves, however, as seen in Figure 18, confirms that for the ASD group, the difference between conditions is more pronounced at anterior electrodes, while for the NT group, the responses show a more posterior distri-

Figure 19: Anterior/Posterior Distribution of Summated P600 effects by Group



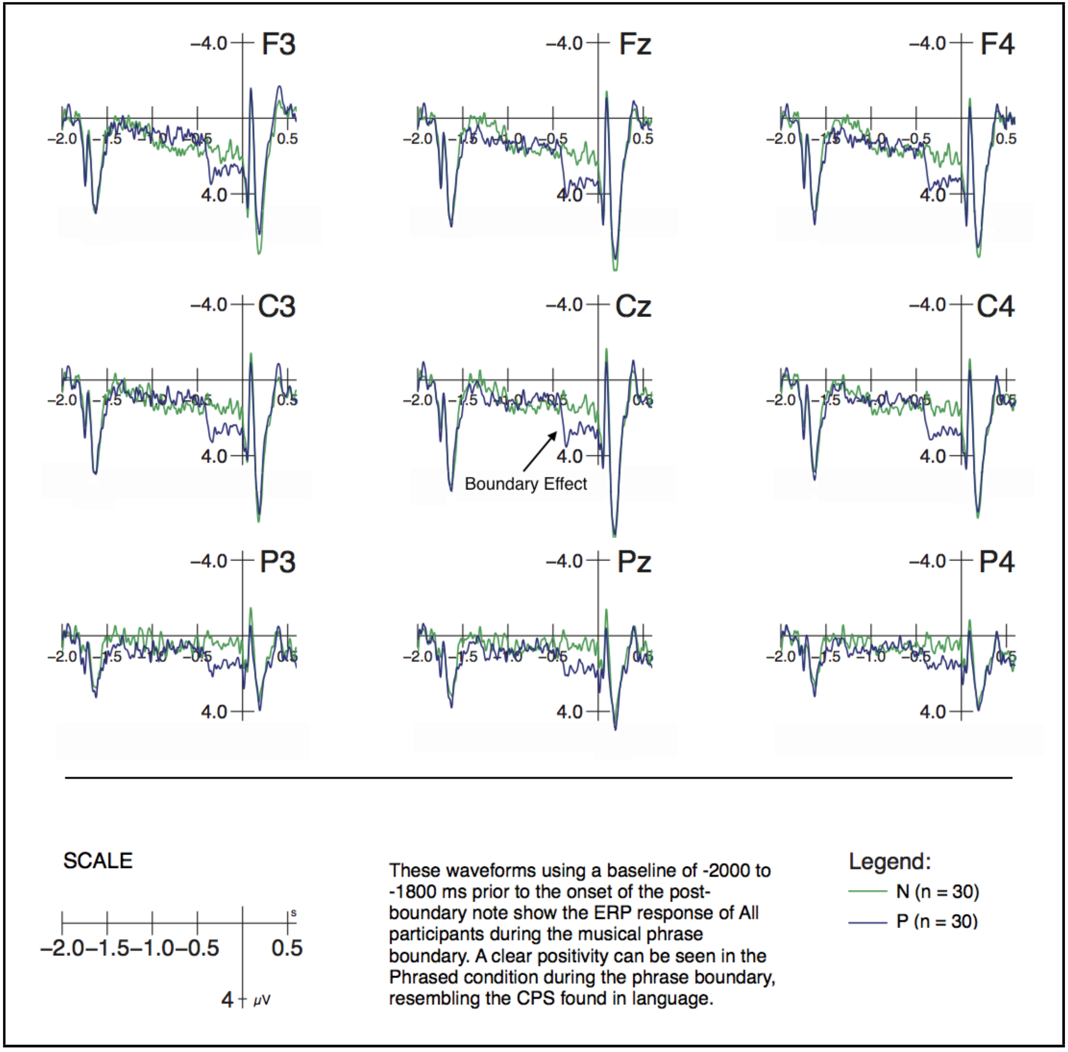
bution. This differential pattern of anterior distribution was also found in CPS1, but there it was solely due to a CPS response. The presence of a more anterior distribution even when CPS effects are subtracted out suggests that P600 effects also demonstrate a more frontal distribution for the ASD group. The frontal offset N1 negativity observed in the ASD group in condition A further heightens the anterior/posterior contrast between these two conditions. This contrast is shown in the anterior/posterior distribution of differences between conditions A and C as shown in Figure 19. While the posterior electrodes demonstrate the greatest additive effect of P600 effects in the NT group, the additive effect of the P600 and CPS is more pronounced at anterior electrodes for the ASD group.

3.2.2 Music

The results of the music portion of the experiment are shown in Figures 20 through 26, while the results of statistical comparisons of music responses are summarized in Tables 12 through 15. All TWs are stated, following convention, in ms relative to the onset of the first note following the phrase boundary in conditions Phrased (P) and NoPause (N), and the corresponding note (due to relative lack of phrase boundary) in condition UnPhrased (U). For all baselines and TWs, ANOVAs were conducted using the factors Cond (Phrased x UnPhrased x NoPause), Grp (ASD x NT), Hem (L x R), and AntPost (Ant x Mid x Post).

-2000 to -1800 ms Baseline

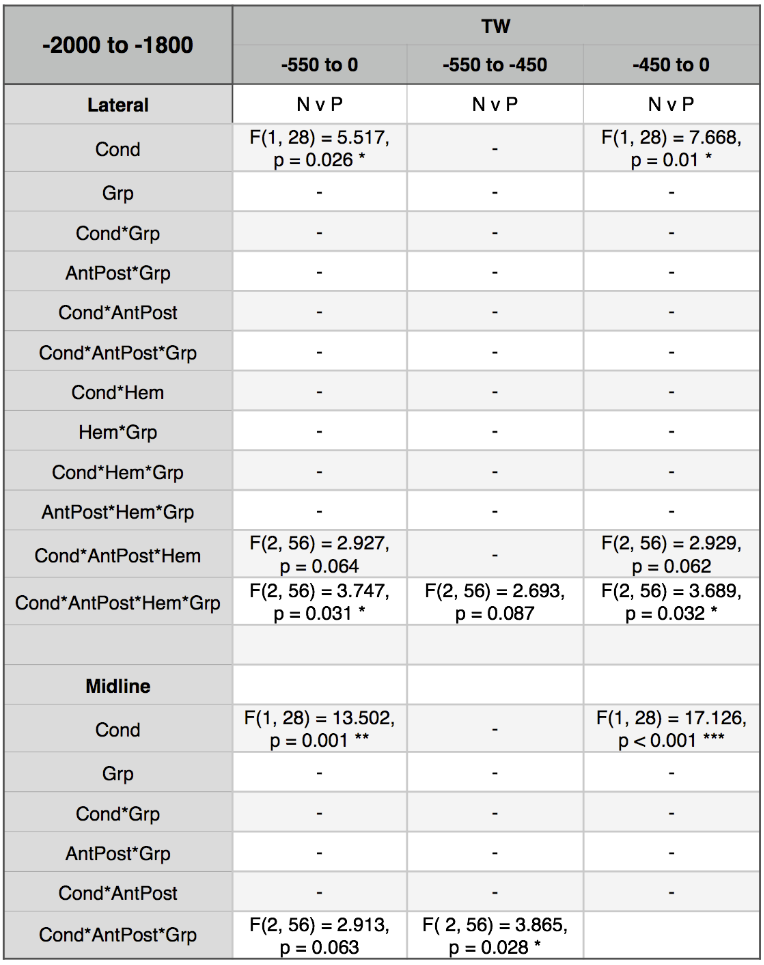
Figure 20: All Participants Baseline -2000 to -1800: Conditions P v N



These waveforms using a baseline of -2000 to -1800 ms prior to the onset of the post-boundary note show the ERP response of All participants during the musical phrase boundary. A clear positivity can be seen in the Phrased condition during the phrase boundary, resembling the CPS found in language.

The first baseline to be discussed is the baseline taken from the 200 ms prior to the final note of the phrase, or -2000 to -1800 ms relative to the onset of the post-boundary note. This period was chosen as it is the last period during which all three conditions are identical prior to the phrase-final note. As can be seen from the comparison of conditions P and N in Figure 20, a sustained positivity during the phrase boundary can be seen in response to condition P. This is the period during which the three conditions differ, as in one condition (U) there were notes, while in the other two, the note was held (N), or rapidly decayed (P). The ERP plots of these two conditions for all participants averaged together show a positivity during the phrase-boundary in response to condition P, which contained a pause. The resulting positivity resembles the CPS found in language, as it occurs during the pause, yet it differs in that it is more frontally maximal for both groups. A TW of 550 ms was analyzed as well as two smaller TWs consisting of -550 to -450 ms, and -450 to 0 ms relative to the offset of the musical phrase boundary due to a sharp positive spike at the beginning of the sustained positivity.

Table 12: Pairwise ANOVAs: -2000 to -1800 Baseline: Conditions N x P



-2000 to -1800	**TW**		
	-550 to 0	**-550 to -450**	**-450 to 0**
Lateral	N v P	N v P	N v P
Cond	F(1, 28) = 5.517, p = 0.026 *	-	F(1, 28) = 7.668, p = 0.01 *
Grp	-	-	-
Cond*Grp	-	-	-
AntPost*Grp	-	-	-
Cond*AntPost	-	-	-
Cond*AntPost*Grp	-	-	-
Cond*Hem	-	-	-
Hem*Grp	-	-	-
Cond*Hem*Grp	-	-	-
AntPost*Hem*Grp	-	-	-
Cond*AntPost*Hem	F(2, 56) = 2.927, p = 0.064	-	F(2, 56) = 2.929, p = 0.062
Cond*AntPost*Hem*Grp	F(2, 56) = 3.747, p = 0.031 *	F(2, 56) = 2.693, p = 0.087	F(2, 56) = 3.689, p = 0.032 *
Midline			
Cond	F(1, 28) = 13.502, p = 0.001 **	-	F(1, 28) = 17.126, p < 0.001 ***
Grp	-	-	-
Cond*Grp	-	-	-
AntPost*Grp	-	-	-
Cond*AntPost	-	-	-
Cond*AntPost*Grp	F(2, 56) = 2.913, p = 0.063	F(2, 56) = 3.865, p = 0.028 *	

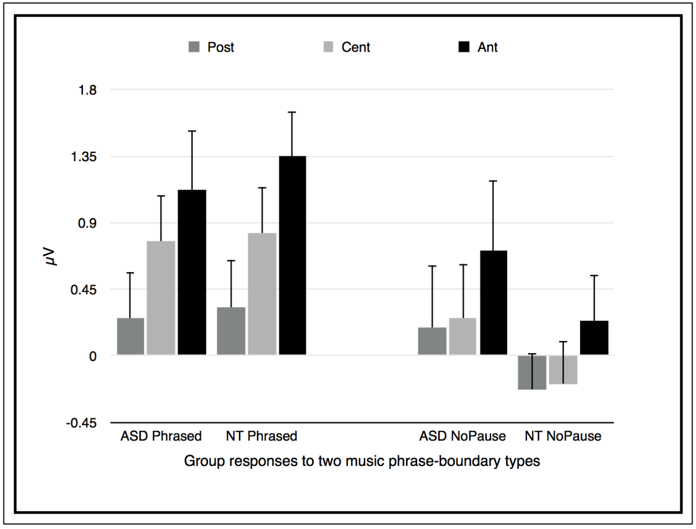
For the global ANOVA for the TW of -550 to 0, there were no significant differences between conditions, with the exception of a trend toward interaction of Cond x AntPost x Hem (P = 0.095). The pairwise comparison of N x P, (Table 12) however, showed a main-effect of Cond at both lateral ($F(1, 28) = 5.517$, $p = 0.026$, $d = 0.268$) and midline ($F(1, 28) = 13.502$, p = 0.001, $d = 0.592$) electrodes. For all participants, the difference between responses to the presence of a rest, or pause, contrasted sharply, especially at midline electrodes, with responses to the phrase-final note being held for the duration of the phrase-boundary. There were no main-effects of Grp, but there was a significant interaction of Cond x AntPost x Hem x Grp ($F(2, 56) = 3.747$, $p = 0.031$) and the interaction of Cond x AntPost x Grp approached significance at midline sites ($p = 0.063$).

The AntPost interactions were driven by a more anterior distribution of effects for the ASD group in response to condition N (NoPause). Figure 21 shows the anterior/posterior distribution of the boundary effect for both groups in response to the two conditions. The ASD group response to the NoPause condition is similar to that observed for condition P (Phrased), yet with a smaller amplitude and more gradual trend toward positivity, as compared to the sharp onset of the condition P response. The interaction involving the factor Hem was due to a slight left lateralization of the NT response to condition P. For a further exploration of this lateralization effect see the section on -1800 to -600 Baseline results.

There were no significant effects or interactions in the comparison of conditions P x U, but trends toward significance in the pairwise comparison of conditions N x U were observed. The presence of additional notes in only one of these conditions (U), however, makes comparisons with condition U relatively uninformative in this TW.

In the TW of -550 to -450, examined due to the apparent positive spike at the offset of the phrase-final note in condition P, the only comparison that reached significance was the interaction of Cond x AntPost x Grp. Once again, this was due to a similar response in the ASD group to conditions N and P.

Figure 21: Anterior/Posterior Boundary Effect Responses for Conditions N and P



In the -450 to 0 TW, which was the period during which one condition contained a rest, no results of global ANOVAs reached significance, yet in the pairwise comparison of N x P, there were significant main-effects of Cond at both lateral (F(1, 28) = 7.668, p = 0.01, $d = 0.327$) and midline (F(1, 28) = 17.126, p < 0.001, 0.687) sites. These results show that the strongest phrase-boundary response occurs at midline electrodes, and after the offset of the phrase-final note, at which point the positivity almost immediately begins and is sustained for the duration of the phrase-boundary. Once again, the only interaction involving the factor Grp was Cond x AntPost x Hem x Grp ($F(2,\ 56) = 3.689$, $p = 0.032$) due to the differences in condition N responses noted above. No other results of the pairwise comparisons or global ANOVAs reached significance at the level of $p < 0.05$.

Taken together, these results suggest that the TW of -450 to 0 contains the strongest differences between responses to the two conditions, confirming the hypothesis that a pause during the phrase boundary elicits a positive shift. The responses vary between groups, however, as the ASD group shows similar responses to both Phrased and NoPause conditions, as opposed to the NT group whose response to the NT condition is negligible. Further, the fact that no significant main-effects of Cond were observed in the TW of -550 to -450, while there were significant effects between -450 and 0 suggests that the positive ERP response is not simply due to an offset positivity resembling a P2, but is a sustained positivity resembling the language CPS.

This same baseline was also used to examine the TW corresponding to the previously described 'music CPS'. This TW occurs following the phrase-boundary as opposed to during it, and is shorter in duration than the CPS observed in response to language. No main-effects of condition were observed for this TW using this baseline, and the only interaction that reached significance at the level of $p < 0.05$ was a Cond x Grp interaction comparing conditions N x U ($F(1, 28) = 5.768$, $p = 0.023$), likely due, once again, to different responses to condition N. This TW was also examined using the more proximate baseline of -200 to 0 ms, which will be discussed in the next section.

The TWs of -1600 to -1300, and -1200 to -600 relative to the phrase-boundary offset were chosen for analysis based on visual inspection of ERP plots. These TWs correspond to +200 to +500 ms, and +600 to +1200 ms after the onset of the phrase-final note, respectively.

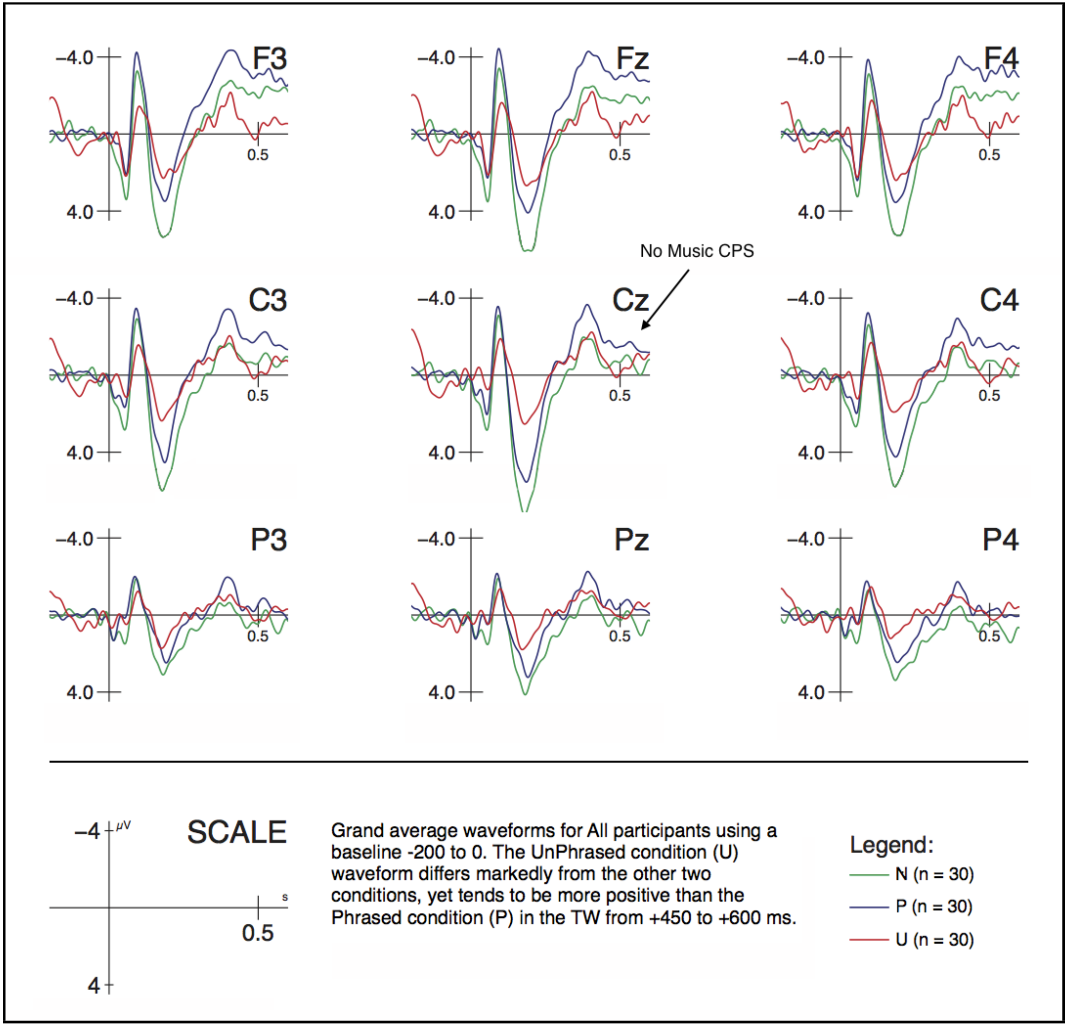
In the first TW from -1600 to -1300, there was a main-effect of Cond at midline electrodes ($F(1, 28) = 5.669$, $p = 0.024$, $d = -0.478$) as well as several significant interactions, nearly all involving the factor Grp. There was a significant main-effect of Grp ($p = 0.046$) in the P x U pairwise comparison, as well as an interaction of AntPost x Grp ($p = 0.043$), suggesting that the effect was more anterior, and only present in the ASD group. This period follows the onset of the phrase-final note in all conditions, which are all nearly identical. This effect could be due to heightened lower-level auditory processing responses following the phrase-final notes, and appears to be confined only to condition U. Further investigation is needed to explain the origin of this between-group difference. No comparisons reached significance at the level of $p < 0.05$ in the TW of -1200 to -600 ms.

Baseline -200 to 0

The baseline -200 to 0 was only used to examine the previously described music CPS in the TW of +450 to +600 ms relative to the post-boundary onset in order to replicate previous studies (see Methods). As can be seen from the grand average waveforms of all three conditions (Figure 22), condition U responses tended to be smaller in general, likely due to the preceding context of notes, leading to smaller N1/P2 components. Further, during the TW under examination, response to condition U tended to be more positive than the other two conditions, particularly at frontal and lateral sites. While during this TW condition N did become briefly more positive than condition U, condition P remained either more negative relative to U, or approximately the same, contrary to previous findings predicting a positivity in response to the Phrased (P) condition.

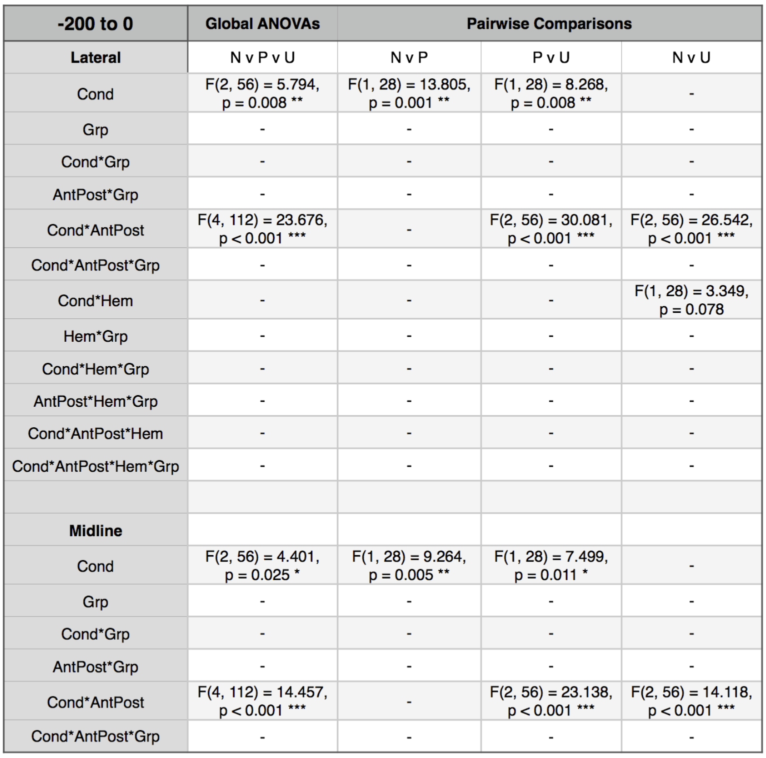
The global and pairwise comparisons are shown in Table 13. For the ANOVAs using the three levels of factor Cond (P x N x U), there was a significant main-effect of Cond at both lateral ($F(2, 56) = 5.794$, $p = 0.008$) and midline ($F(2, 56) = 4.401$, $p = 0.025$) sites. These differences are due to a generally more positive response in the UnPhrased condition, as opposed to the predicted positivity in response to the Phrased condition. The negativity in condition P is almost certainly due simply to the placement of the baseline during the phrase-boundary where responses to this condition show a sustained, anteriorly maximal positivity. This interpretation is supported by the highly significant interaction of Cond x AntPost for both lateral ($F(4, 112) = 23.676$, $p < 0.001$) and midline ($F(4, 112) = 14.457$, $p < 0.001$) analyses. The pairwise comparison between these two conditions is nearly identical to that shown in the global ANOVA, showing that, indeed, these results are driven by the comparison of conditions P and U. The statistical results confirm what the waveforms show, notably that the greatest differences between conditions N/P vs U are frontally distributed. Once again, these differences are almost certainly due to the presence of a positivity during the baseline in the Phrased condition.

Figure 22: All Participants Baseline -200 to 0: Conditions P v N v U



Grand average waveforms for All participants using a baseline -200 to 0. The UnPhrased condition (U) waveform differs markedly from the other two conditions, yet tends to be more positive than the Phrased condition (P) in the TW from +450 to +600 ms.

Table 13: ANOVAs: -200 to 0 Baseline



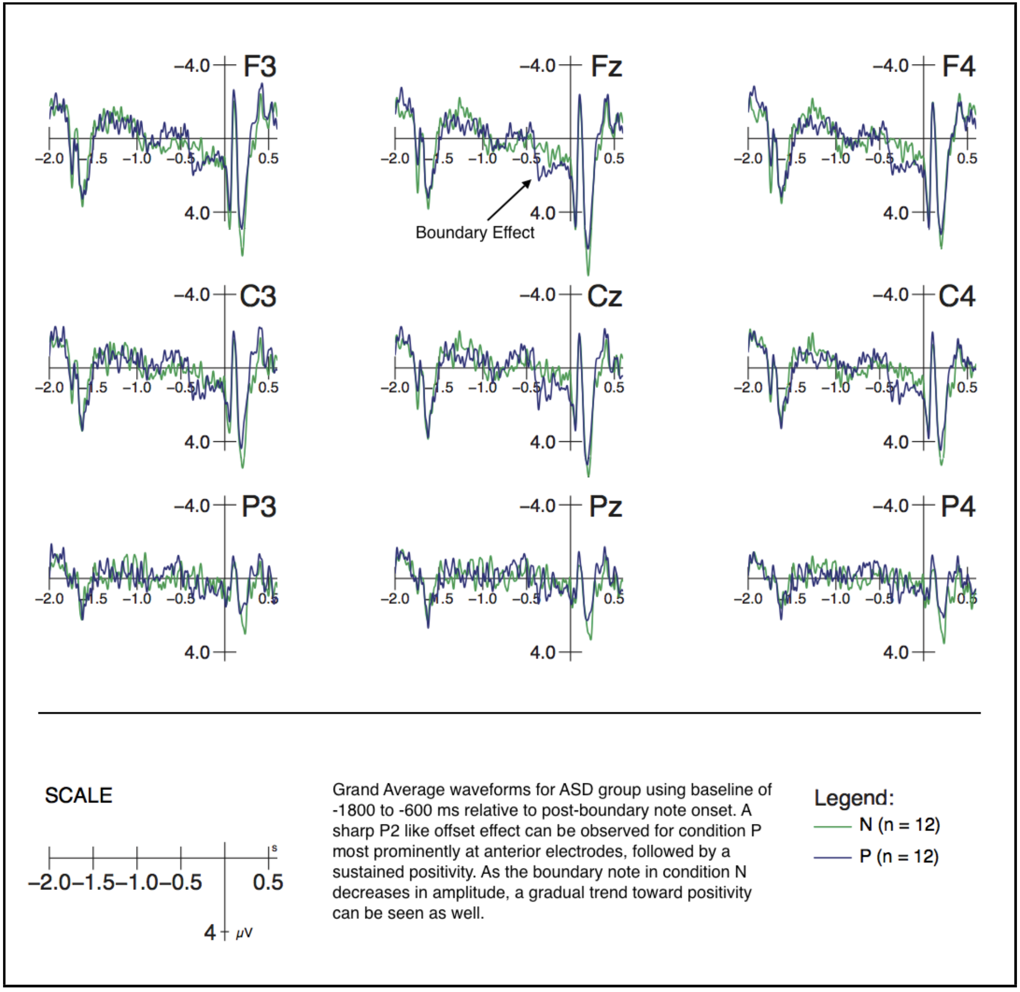
-200 to 0	**Global ANOVAs**	**Pairwise Comparisons**		
Lateral	N v P v U	N v P	P v U	N v U
Cond	F(2, 56) = 5.794, p = 0.008 **	F(1, 28) = 13.805, p = 0.001 **	F(1, 28) = 8.268, p = 0.008 **	-
Grp	-	-	-	-
Cond*Grp	-	-	-	-
AntPost*Grp	-	-	-	-
Cond*AntPost	F(4, 112) = 23.676, p < 0.001 ***	-	F(2, 56) = 30.081, p < 0.001 ***	F(2, 56) = 26.542, p < 0.001 ***
Cond*AntPost*Grp	-	-	-	-
Cond*Hem	-	-	-	F(1, 28) = 3.349, p = 0.078
Hem*Grp	-	-	-	-
Cond*Hem*Grp	-	-	-	-
AntPost*Hem*Grp	-	-	-	-
Cond*AntPost*Hem	-	-	-	-
Cond*AntPost*Hem*Grp	-	-	-	-
Midline				
Cond	F(2, 56) = 4.401, p = 0.025 *	F(1, 28) = 9.264, p = 0.005 **	F(1, 28) = 7.499, p = 0.011 *	-
Grp	-	-	-	-
Cond*Grp	-	-	-	-
AntPost*Grp	-	-	-	-
Cond*AntPost	F(4, 112) = 14.457, p < 0.001 ***	-	F(2, 56) = 23.138, p < 0.001 ***	F(2, 56) = 14.118, p < 0.001 ***
Cond*AntPost*Grp	-	-	-	-

Baseline -1800 to -600

The baseline of -1800 to -600 was chosen to examine phrase-boundary effects using a more temporally proximate baseline than the earlier correction of-2000 to -1800. For this closer baseline, the effects found using the other baseline were generally confirmed. As can be seen from Figures 23 and 24, both groups demonstrate a sustained positivity in response to the Phrased condition, while the ASD group also shows a similar response with a slightly delayed onset latency to the NoPause condition.

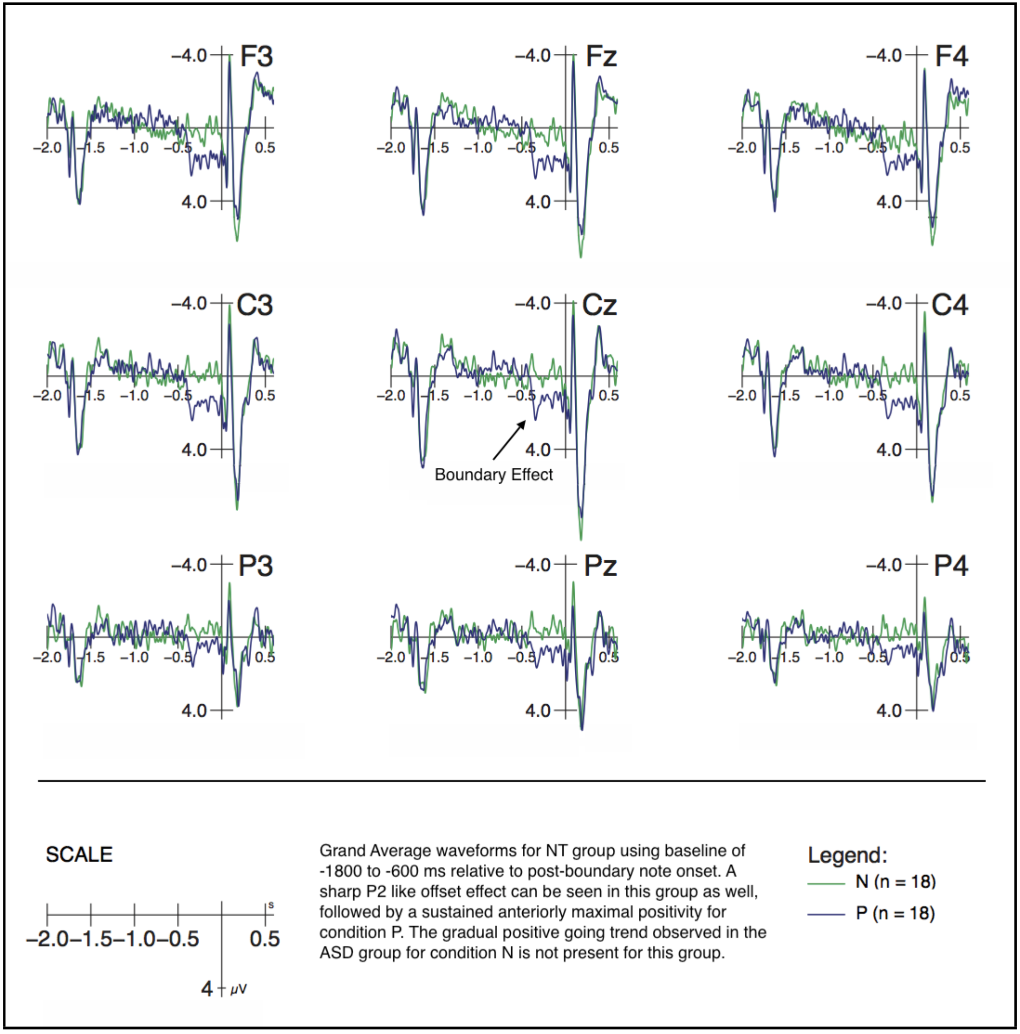
A main-effect of Cond was observed during the TW -550 to 0 in the global ANOVAs at both lateral and midline sites both at the significance level of $p < 0.05$. (Table 14) Additionally, Cond x AntPost interactions were observed for both analyses (lateral: $F(4, 112) = 3.315$, $p = 0.031$; midline: $F(4, 112) = 4.207$, $p =$

Figure 23: ASD Baseline -1800 to -600: Conditions P v N



Grand Average waveforms for ASD group using baseline of -1800 to -600 ms relative to post-boundary note onset. A sharp P2 like offset effect can be observed for condition P most prominently at anterior electrodes, followed by a sustained positivity. As the boundary note in condition N decreases in amplitude, a gradual trend toward positivity can be seen as well.

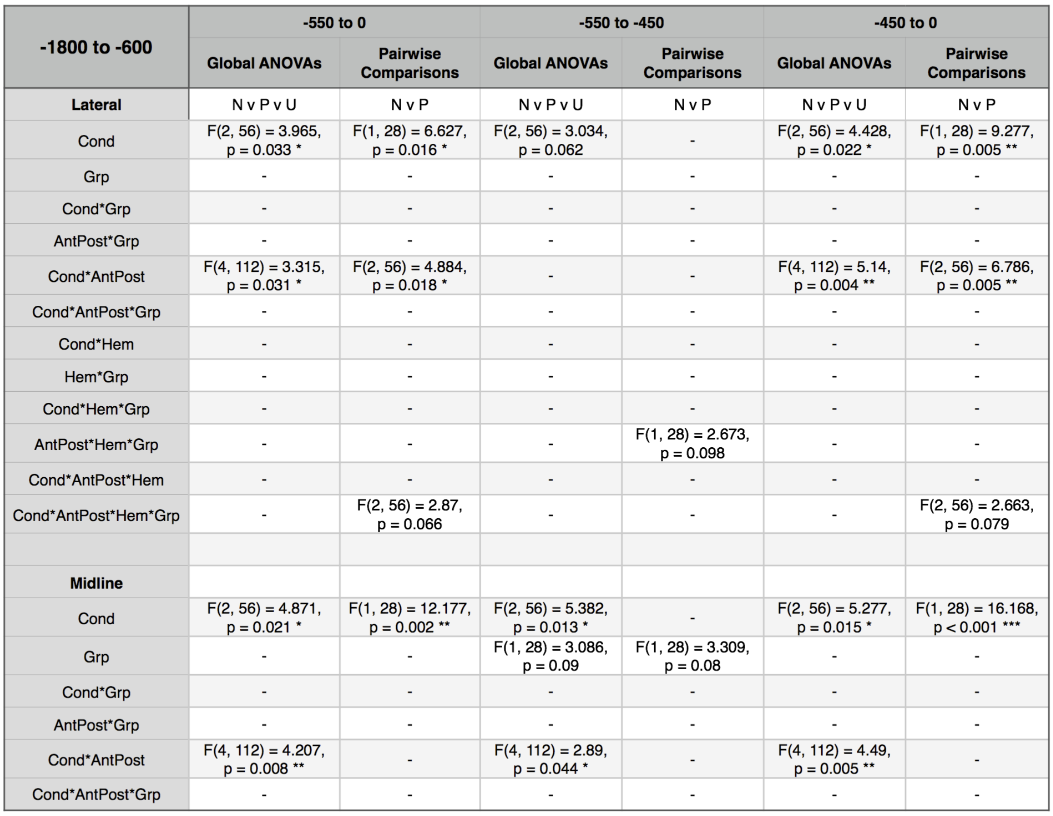
Figure 24: NT Baseline -1800 to -600: Conditions P v N



Grand Average waveforms for NT group using baseline of -1800 to -600 ms relative to post-boundary note onset. A sharp P2 like offset effect can be seen in this group as well, followed by a sustained anteriorly maximal positivity for condition P. The gradual positive going trend observed in the ASD group for condition N is not present for this group.

0.008). While the effects of Cond were also seen in the pairwise comparison of N x P (lateral: $F(1, 28) = 6.627$, $p = 0.016$, $d = 0.42$); midline: $F(1, 28) = 12.177$, $p = 0.002$, $d = 0.644$), the Cond x AntPost interaction was observed only at lateral electrodes between these two conditions ($F(2, 56) = 4.884$, $p = 0.018$). These patterns of significance confirm those found using the earlier baseline, showing that the presence of a pause or rest during the phrase-boundary elicits an anteriorly maximal positivity similar to that found in language. No interactions involving the between-subject factor Grp reached significance at the level of $p < 0.05$, however, a trend toward significance was again observed for the interaction of Cond x AntPost x Hem x Grp in the pairwise comparison of N x P ($F(2, 56) = 2.87$, $p = 0.066$) due, once again, to the more anterior response of the ASD group to condition N.

Table 14: ANOVAs: -1800 to -600 Baseline: Conditions N x P x U

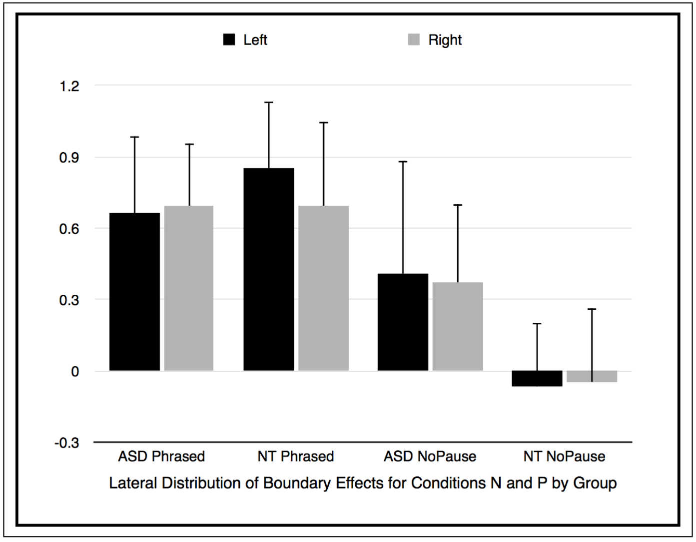


-1800 to -600	**-550 to 0**		**-550 to -450**		**-450 to 0**	
	Global ANOVAs	**Pairwise Comparisons**	**Global ANOVAs**	**Pairwise Comparisons**	**Global ANOVAs**	**Pairwise Comparisons**
Lateral	N v P v U	N v P	N v P v U	N v P	N v P v U	N v P
Cond	F(2, 56) = 3.965, p = 0.033 *	F(1, 28) = 6.627, p = 0.016 *	F(2, 56) = 3.034, p = 0.062	-	F(2, 56) = 4.428, p = 0.022 *	F(1, 28) = 9.277, p = 0.005 **
Grp	-	-	-	-	-	-
Cond*Grp	-	-	-	-	-	-
AntPost*Grp	-	-	-	-	-	-
Cond*AntPost	F(4, 112) = 3.315, p = 0.031 *	F(2, 56) = 4.884, p = 0.018 *	-	-	F(4, 112) = 5.14, p = 0.004 **	F(2, 56) = 6.786, p = 0.005 **
Cond*AntPost*Grp	-	-	-	-	-	-
Cond*Hem	-	-	-	-	-	-
Hem*Grp	-	-	-	-	-	-
Cond*Hem*Grp	-	-	-	-	-	-
AntPost*Hem*Grp	-	-	-	F(1, 28) = 2.673, p = 0.098	-	-
Cond*AntPost*Hem	-	-	-	-	-	-
Cond*AntPost*Hem*Grp	-	F(2, 56) = 2.87, p = 0.066	-	-	-	F(2, 56) = 2.663, p = 0.079
Midline						
Cond	F(2, 56) = 4.871, p = 0.021 *	F(1, 28) = 12.177, p = 0.002 **	F(2, 56) = 5.382, p = 0.013 *	-	F(2, 56) = 5.277, p = 0.015 *	F(1, 28) = 16.168, p < 0.001 ***
Grp	-	-	F(1, 28) = 3.086, p = 0.09	F(1, 28) = 3.309, p = 0.08	-	-
Cond*Grp	-	-	-	-	-	-
AntPost*Grp	-	-	-	-	-	-
Cond*AntPost	F(4, 112) = 4.207, p = 0.008 **	-	F(4, 112) = 2.89, p = 0.044 *	-	F(4, 112) = 4.49, p = 0.005 **	-
Cond*AntPost*Grp	-	-	-	-	-	-

Breaking this larger TW into the smaller TWs demonstrate the origin of this Grp interaction. In the TW of -550 to -450 ms there were trends toward significance for the main-effect of Grp at midline electrodes ($F(1,\ 28) = 3.309$, $p = 0.08$), and the interaction of AntPost x Hem x Grp ($F(1,\ 28) = 2.673$, $p = 0.098$) in the pairwise comparison of N x P. Thus, it appears that the ASD group shows a slightly greater sensitivity to the onset of the phrase-boundary in both conditions.

In the TW of -450 to 0, there were again more significant effects than in the window from -550 to 0. While main-effects of Cond in the global ANOVAs remain significant at the level of $p < 0.05$, the Cond x AntPost interaction increases in significance ($F(4,\ 112) = 5.14$, $p = 0.004$). Further, in the pairwise comparisons of N x P, the main-effect of Cond increases in significance for both lateral ($F(1,\ 28) = 9.277$, $p = 0.005$, $d = 0.502$), and midline ($F(1,\ 28) = 16.168$, $p < 0.001$, $d = 0.759$) comparisons. Once again the trend toward a significant interaction of Cond x AntPost x Hem x Grp appears ($F(2,\ 56) = 2.663$, $p = 0.079$). An exploration of this effect is shown in Figure 25, where both groups' responses to conditions N and P are shown by hemisphere. This chart shows that for the ASD group, both hemispheres show larger positive responses to the NoPause condition than the NT group, as well as showing that for the NT group, the boundary effect to the Phrased condition is slightly left-lateralized.

Figure 25: Hemispheric Comparisons of Boundary Effect for Conditions N and P



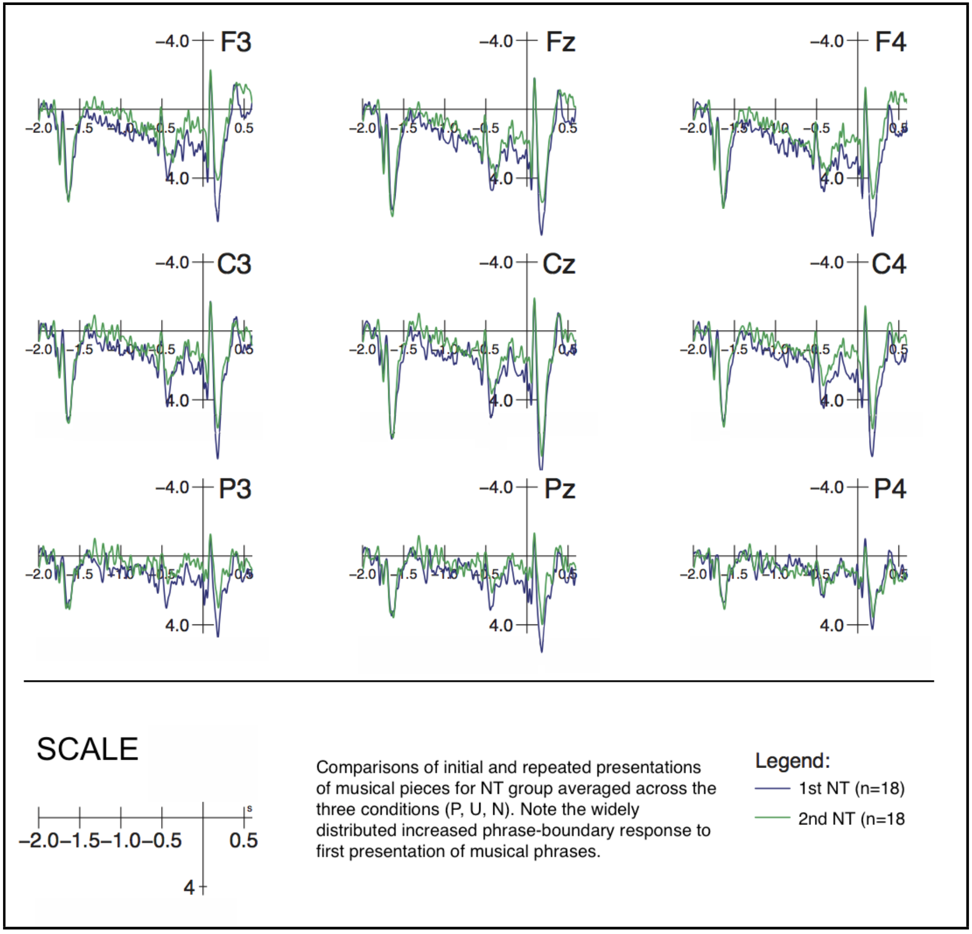
Using this temporally closer baseline results in starker differences between responses to phrase-boundary type. Between group differences are largely confined to the TW corresponding to the onset of the phrase-boundary, and diminish during the actual boundary, with the exception of distribution differences. These differences result from a more positive response to the NoPause condition in the ASD group, suggesting heightened processing of phrase-boundaries regardless of the presence of a pause. Also, the ASD group's response to condition P was slightly more right lateralized as opposed to the response of the NT group, who showed a larger left lateral response.

Cadence & Repetition Effects

Both cadence and repetition effects were examined using the baseline from -2000 to -1800 ms relative to the post-boundary note. Comparisons involving cadence effects contrasted responses to the first and third presentation of the melodies which contained a half-cadence ending on the third or fifth scale degree, (harmonically the tonic chord), versus the second and fourth presentation of the melodies which ended with a full-cadence on the first scale degree (also the tonic). These comparisons were averaged across all three conditions U, N, and P, and thus had more power than the three condition comparisons mentioned earlier. No effects of cadence reached significance in any of the TWs examined except for the TW between +450 to +600. In this TW, however, there was an imbalance of notes across conditions resulting from the creation of musical stimuli that follow the rules of western musical composition. The post-boundary TW contained a one-(9/12) or two-note anacrusis (3/12), but the final post-boundary notes of each piece always consisted of a single note. Thus comparisons in this TW are obfuscated by qualities of the stimuli. Additionally, despite a slight observed anterior positivity between -1200 and -600 ms, or 600 ms following the onset of the phrase-final note in the half-cadence condition, this effect did not reach significance. Further research is needed to determine the influence of Cadence on electrophysiology.

In terms of repetition effects, it appears from the grand average waveforms (Figures 26 & 27) that the phrase-boundary effect is strongest at anterior sites for earlier presentations of stimuli for both groups. This difference resulted in a significant main-effect of Cond in the TWs from both -550 to 0 ($F(1, 28) = 4.73$, $p = 0.038$, $d = -0.401$) and -450 to 0 ($F(1, 28) = 4.41$, $p = 0.045$, $d = -0.382$) at midline electrodes. More research still needs to be conducted in order to determine the extent of this progressively diminishing effect, yet it may be due to a decreased need to consolidate phrase-boundaries and create expectations about what is to follow. As the musical phrases were repeated almost exactly between the first presentation and the repeat, participants were likely quite quickly able to learn that the second presentation was identical. This was particularly the case

Figure 26: NT Participants: Initial vs. Repeated Presentation of Musical Phrases

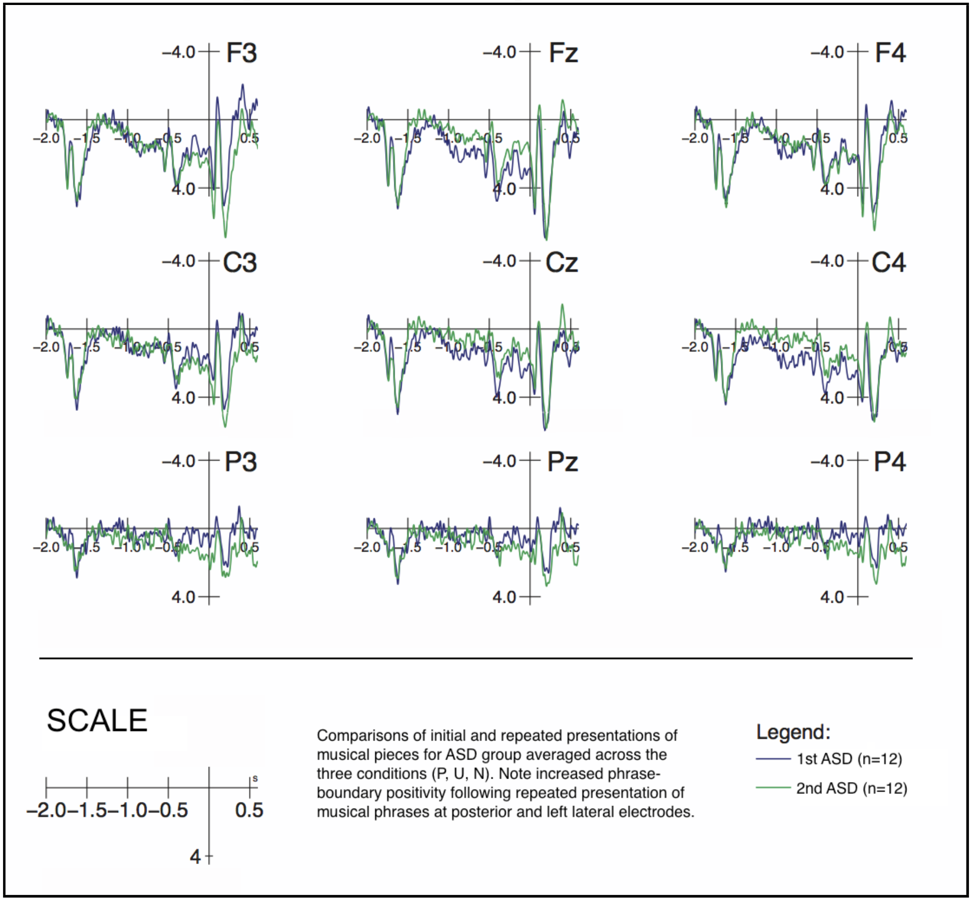


Comparisons of initial and repeated presentations of musical pieces for NT group averaged across the three conditions (P, U, N). Note the widely distributed increased phrase-boundary response to first presentation of musical phrases.

at posterior electrodes, which resulted in a Cond x AntPost interaction at midline sites (F(2, 56) = 3.824, p = 0.035).

For the ASD group, the effect of repetition on responses was reversed, particularly at lateral and posterior electrodes, resulting in Cond x Grp interactions for both TWs between -550 to 0 (F(1, 28) = 6.569, p = 0.016) and -450 to 0 (F(1, 28) = 7.166, p = 0.012) at the lateral electrodes. In fact, at posterior electrodes for the ASD group, the second presentation of the phrase boundary resulted in a relatively larger, more positive response, contrary to the diminished response for the NT group, resulting in a significant Cond x AntPost x Grp interaction at midline electrodes between -550 and -450 ms (F(2, 56) = 3.591, p = 0.047). This suggests that the ASD group may have been able to more accurately process the phrase-boundary on second and later presentations. Considering that repetitive

Figure 27: ASD Participants: Initial vs. Repeated Presentation of Musical Phrases



Comparisons of initial and repeated presentations of musical pieces for ASD group averaged across the three conditions (P, U, N). Note increased phrase-boundary positivity following repeated presentation of musical phrases at posterior and left lateral electrodes.

behavior is one of the diagnostic criteria for ASDs, and that anecdotally, many ASD individuals prefer routine, perhaps the effects of repetition allow this group to focus on higher-level, or global features of input.

4 Discussion

4.1 Language

4.1.1 Behavioral Response

In terms of behavioral response, both the NT and ASD groups rated condition C as the least natural, suggesting that the violation created by splicing was perceived to some extent by the majority of participants in both groups. Further, both the NT group and All participants averaged together rated condition B as less natural than condition A. A possible reason for this could have been the long, consistent, silenced pause at the first IPh of this sentence. Any number of elements, such as the consistency of this length, the length itself, the presence of silence, or indeed some other factor altogether could have caused this lower rating of naturalness. Regardless, the fact that there is no Cond x Grp interaction between the ASD group and NT group's responses suggests that the ASD group was successfully able to perceive the violation present in condition C, even if their rating of condition B as less natural than condition A did not reach significance. Thus, the incongruous syntax/prosody interaction was perceived as such.

That the ASD group consistently perceived this violation was somewhat contrary to expectations. The premise of the experiment was based on the dual-pathway processing model developed by Friederici & Alter. (2004) This theory states that prosodic information, particularly on the global level, is processed in a distinct pathway, which is largely localized in the right-hemisphere. Syntactic information, on the other hand, is processed largely in the classical language areas in the left hemisphere, including, but not limited to, Broca's area in the IFG. The ability to reconciling these two pathways provides counter-evidence to both the theories of under-connectivity and weak central coherence (WCC). Further, individuals with ASDs have repeatedly been found to be deficient in receptive prosodic abilities. (i.e. Paul, et al., 2005) Evidence from the behavioral portion of this experiment does not support this previous observation, as behaviorally, the ASD participants seemed to note the violation in the incongruous prosody/syntax condition.

The ability to perceive this type of violation is likely far from universal in individuals with ASDs. Previous research has noted a pattern of linguistic delay, rather than deviance, as mentioned earlier. Perhaps the fact that the ASD participants in this study were adults means that they had overcome this delay. Also, the participants in this study were quite high-functioning, as compared to a general ASD population. All participants were verbal, and capable of maintaining conversation with the researchers during application of the EEG cap, and following the experiment. Some participants (through informal self-report) had earned an advanced degree, started a business, or learned to speak a foreign language. Indeed, the verbal-IQs of the ASD group were high, even compared to the NT group (but not significantly different, see Table 1). Many had also developed social coping strategies for being in unfamiliar surroundings, including informing themselves about the background of the researchers in order to have subject matter for conversation. In addition to the offered compensation for participation (finding regular employment remained problematic for some due to difficulties with the interview process), some participants also expressed a desire to participate in order to increase understanding of individuals with ASDs, and to potentially improve the lives of those diagnosed in the future. Just the fact that the individuals in this study contacted the researchers independently through phone or email, and secured their own transportation to and from the study site demonstrates the relatively high-functioning nature of the ASD individuals who participated in this study.

Despite the high verbal and non-verbal abilities of the participants, many also mentioned irony, or emotional content as being the most difficult aspects of language for them to understand. Whether this was due to familiarity with trends related to their own diagnosis, and thus hoping to demonstrate qualifications for the study, or legitimate problems related to difficult experiences is uncertain. It can be expected, however, that a similar experiment conducted with a lower-functioning sample, perhaps recruited solely through psychiatrist recommendation, would show results that differ more greatly between NT and ASD groups.

4.1.2 EEG

Closure Positive Shift (CPS)

By using the neuro-typical population, this study sought to confirm the results of previous CPS studies using similar paradigms. In so doing, this study had a control from which to examine variations in electrophysiological response in a population with ASDs. This purpose of the study was accomplished, as this study confirmed the results from previous CPS studies.

In the first TW examined, a CPS was clearly present for the NT group in response to condition B, which contained an Intonational Phrase boundary (IPh), as compared to condition A, which did not. The CPS was elicited almost immediately following the onset of the pause, likely due to pre-final phrase boundary characteristics, such as intonational contour change, and a lengthened final syllable. These characteristics have been observed repeatedly to be characteristic of phrase-final prosody, both prior to the first CPS studies (Selkirk, 1984), and in many previous CPS studies conducted. (Steinhauer, et al., 1999; Pannekamp, et al., 2005)

The positive shift resulting from the IPh was defined by a broad distribution that was centro-parietally maximal lasting between 300 and 500 ms, nearly the entire length of the pause. It reached its maximal amplitude between 150 and 200 ms at electrodes Cz and Pz, after which it decreased in positivity, as well as in its widespread distribution until approximately 500 ms. This effect was also maximal over the right-hemisphere, which is the hemisphere largely responsible for processing global elements of prosody and many facets of music. The right-hemispheric maximality for this effect suggests that it is indeed a prosodic response. Additionally, the prolonged nature of this positivity clearly distinguishes it from lower-level acoustic/perceptual ERP components such as a P200, which has been described following the offset of acoustic stimuli. Further, that the CPS only develops in children with the acquisition of syntax (contrary to the P200, which appears much earlier: Männel & Friederici, 2009; 2011), confirms that it represents higher-level prosodic processing. (see Steinhauer (2003) for a more detailed argument)

The interpretation of this effect is that it serves as a prosodically triggered chunking mechanism responsible for consolidating the information in the previous phrase, and identifying phrase boundaries to aid in hierarchical component structuring to effectively parse the sentence. Whether this component includes an anticipatory element is unclear, but that it signals the 'closure' of phrases, both at the end of clauses, as well as sentence-finally, is well established. (Roll, et al., 2011)

CPS in ASD Group

The ASD group appears to show a slight CPS in response to the IPh in condition B as compared to condition A, which can be seen in grand average waveforms of the EEG data. This relative positive response, however, showed a diminished amplitude, and a more frontal distribution, as compared to the NT group. Also, while the NT group show effects of this positivity over both hemispheres, the ASD group's response was almost entirely limited to right-hemisphere electrodes.

Statistical analysis of electrophysiological responses to these conditions showed significant effects of between-group comparisons, both as a result of condition, and with interactions involving distribution effects. As mentioned earlier, a smaller sample of the NT group was also analyzed and the results were virtually identical to those of the larger group, showing that sufficient trials and participants were available to provide a similar response. Additionally, by examining the difference waves between the two conditions for the individual groups, it can be seen that the effect for the ASD group is smaller, has a different distribution (more anterior, and confined to the right-hemisphere) and while it appears that this effect has a shorter duration, this difference was not significant. Together these results strongly suggest that even if the ASD group can appropriately perceive the prosodic cues triggering phrase boundary processing responses, these are consolidated in a different manner in individuals with ASDs.

The presence of an observably different response characterized by diminished amplitude and apparent shorter duration suggests that individuals with ASDs employ different neurological IPh processing strategies. While these individuals may be more sensitive to the raw acoustic features of prosody, the prosodic cues do not appear to trigger a chunking mechanism as robust as that found in neurotypicals. Further, the degree of variation within the ASD group, as represented in the error bars of graphs, as well as inspection of the individual averaged waveforms, shows that while the CPS is present in some individuals, it is not present in all. The well-established heterogeneity in this disorder, as well as the number of left-handed or ambidextrous individuals present in this sample likely contributes to the high degree of observed variation. Reversed or absent neural asymmetry has been commonly identified in ASD individuals (i.e., De Fossé, et al., 2004), which could also lead to a type of phase cancellation in ERP responses averaged across individuals with opposite, or dissimilar dipole orientation, causing a reduction in observed electrophysiological effects.

Part of the challenge with interpreting the CPS response in the ASD participant pool is due to a large negative going spike occurring roughly 75-150 ms following the onset of the pause. This spike may be due to an N100 like offset effect driven by the transition from the presentation of acoustically rich spoken language to silence. (Pratt, et al., 2005; 2007) This response, while not observed in the present study in NT populations, has been repeatedly observed in response to the offset of auditory input, however, generally not to language. This potential is referred to as the composite N1 component, consisting of two distinct negativities, the N1a, an automatic auditory response, and the N1b an auditory change detection response distinct from the MMN. (Pratt, 2005) A heightened sensitivity to auditory input in the ASD group may well be responsible for this peak, obscuring the clear positive shift of the CPS by reducing its overall amplitude. The spike is larger than the surrounding alpha waves, suggesting that it represents a significant neurophysiological response. That a nearly identical negativity can be observed in the second CPS supports this interpretation.

The diminished, more anterior CPS response of the ASD group combined with this auditory evoked negativity shows that this group demonstrates substantially different responses to prosodically cued information as compared to that seen in neuro-typicals. In order to identify the neural structures involved in this differential prosodic processing, additional source-localization is necessary. However, the fact that the response was more anterior and largely confined to the right-hemisphere suggests that the prosodic processing pathway is contributing to prosodic consolidation. Perhaps the reason for the diminished response is a lack of contribution from left-hemisphere and more posterior areas involved in other aspects of language, and possibly attention. If this is indeed the case, some type of under-connectivity, or bandwidth limitation, may be the cause.

The variability of response and possibility of electrophysiological potential cancelation due to reversed asymmetry also supports an altered and heterogeneous neural organization, consistent with current understanding of ASDs. The interference of a composite N1 like offset effect further supports the well established autistic tendency toward heightened sensitivity to auditory stimuli, or local bias at the expense of the global. In the present experiment, the confounding effects of low-level auditory response presented challenges to a full understanding of the prosodic ability indexed by the CPS.

Overall, the expressive and perceptive prosodic challenges often experienced by individuals with ASDs are supported by this electrophysiological evidence. Whether the altered pattern of responses is due to a physiological anterior-posterior, or left/right hemisphere long-range connection deficit limiting re-integration of the dual language processing pathways remains uncertain. Predictions made with this hypothesis in mind are, however, supported.

Garden Path Effects

A Garden Path effect (N400/P600) was observed in the neuro-typical group following the incongruous syntactic/prosodic splicing in condition C, replicating earlier studies' findings of a bi-phasic Garden Path effect in Part 2 of the sentences. A bi-phasic pattern, however, was found in all conditions, suggesting that the negative and positive going elements consist of a compounding of numerous electrophysiological processing components, including the N400 in condition C, CPS2 in all conditions, and P600 in condition C. The compounding of these various ERP components makes isolating the Garden Path effects more of a challenge, yet the use of peak-to-peak analysis in the present study showed a significant difference between condition C and the other conditions for the NT group.

Among these components is a negativity leading up to the phrase-boundary in all conditions. This negativity corresponds to the ramp-like negativity observed in previous CPS studies such as Pauker et al. (2011) in several conditions. The ramp-like negativity resembles the expectation related CNV, or Contingent Negative Variation, (Walter, et al., 1964) and may be due to an expectancy of the prosodically and syntactically cued second verb. German is a V2 language, meaning that the finite verb of a declarative clause is always phonologically expressed at the second head node of a sentence. (Wenzlaff, et al., 2005) An additional syntactic operation moves the second verb of a verb phrase (VP) to the end of clauses, following any arguments it may take. The CNV may be cued as a result of an expectancy based on this movement operation, yet more research is needed to understand this component's relation to these specific language phenomena.

In response to condition C, the negativity also includes an N400 component, at least for the NT group. In the grand average waveforms for this group, a negative peak is observed beyond the negative portion of this bi-phasic pattern in response to the prosodic/syntactic violation present in the verb, particularly at posterior electrode sites. This N400 was fairly typical in its distribution, as it has been observed to have a centro-parietal maximality, with its negative peak occurring around 400 ms following the onset of the crucial word. (Kutas & Hillyard, 1980; Koelsch, et al., 2004) The semantic content of the sentences in the spliced condition was unexpected, such as **Kevin promises to sleep Anna*, resulting in this response associated with semantic violations. While the primary violation here is a syntactic one, as cued by the prosody, the argument structure also results in sentences with semantic incongruities, as the meaning of such a sentence is unclear, and, to a certain extent, unresolvable.

The same condition resulted in a P600, peaking around 700 ms, also with a posteriorly maximal effect. Given the type of violation present in these sentences, it is hardly surprising that this effect was observed. As previously mentioned, this effect is well documented in similar CPS studies containing an incongruous condition. The presence of these Garden Path effects was confirmed using peak-to-peak analysis of the TWs corresponding to the peaks typically found in N400 and P600 responses. The average distance between these peaks was measured and compared across conditions, resulting in a significant difference at electrode Pz for the neuro-typical group between conditions A and C, which were acoustically identical at this point, differing only in the presence of a violation.

The ASD group, on the other hand, did not demonstrate significant results of the peak-to-peak analysis of these Garden Path effects. While a clear bi-phasic pattern existed for this group, there were no significant differences between peak-distances in conditions A and C. Examining the grand average waveforms for the ASD group also shows no differences, at least for the N400, and responses to the two conditions pattern together. During the negative phase, there even appears to be a slightly more negative peak (not statistically significant) at central electrodes for condition A, where no violation is present. Thus, while the ASD group did rate the sentences in condition C as less natural sounding than their prosodically congruous counterparts, no Garden Path effects were identifiable through peak-to-peak analysis, most likely due to the lack of an N400 in response to these prosodically incongruous sentences. P600 effects will be discussed further in the section on CPS2, where this effect was more readily identified in the ASD group.

Previous studies have observed both N400 (Braeutigam, et al., 2008; Russo, et al., 2012) and P600 (Koolen, et al., 2013; Koolen et al., 2014) in ASD subjects. Russo et al., (2012) report that N400-like responses were observed in ASD subjects between 150 and 300 ms, and were more posteriorly distributed than that observed in typically developing subjects. Braeutigam, et al., (2008), however, report no differences in latency between a similar subject pool in their study, however, with a weakened N400 response at temporal sites. The authors mention, in a quick review of the literature, that ASD responses to N400 vary substantially by task, with some studies finding no N400 effect, and others finding delayed effects. These conflicting results taken together provide a further demonstration of the heterogeneity of ASDs, as well as some sense of the degree to which this disorder has puzzled researchers.

Recent P600 studies (Koolen, et al., 2013; Koolen et al., 2014) on individuals with ASDs have failed to find a P600 response unless the ASD participants were expressly told to focus on the degree of semantic implausibility in the sentences presented. The lack of N400 in the ASD group could be related to the task instructions in the experiment. Participants were not told to focus on 'plausibility', but rather 'naturalness'. This dimension is rather broad, as it is unclear whether focus should be applied to content, prosody, syntax, or some other acoustic feature of the recordings. From these previous studies, it appears that by drawing attention to specific features of language, researchers were able to aid the ASD subjects in processing linguistic input in a more similar way to neuro-typicals. Perhaps the level of complexity found in typical sentence violation paradigms is too task demanding for this group to focus attention on one specific aspect of the stimuli. An inability to filter out irrelevant information may prevent individuals with ASDs from focusing on higher-level features such as grammaticality, prosody, or plausibility, resulting in a 'local bias'. Often, prosody signals many of these features of language in natural speech, yet prosody is a global phenomenon occurring on the phrasal level, requiring listeners to incorporate several levels of acoustic information simultaneously.

Whatever the reason, the lack of N400 seen in the ASD group in the present experiment is further evidence of the linguistic deficits found in ASDs. It remains unclear, however, whether the absence of this neuro-physiological response is due to difficulty integrating the dual processing streams of language, or simply divergent and highly variable responses to language in general.

CPS 2

During the TWs following the second IPh in conditions B and C, and corresponding to the first IPh in condition A, a CPS response was observed for the NT group in response to all three conditions. Previous studies using similar paradigms have found nearly identical results, showing that in response to this phrase boundary, a CPS is consistently elicited. While this second IPh TW provides further evidence of phrase-boundary processing in all conditions, a proper control condition is lacking, due to the presence of a boundary in all three conditions. Still, we can confidently identify that a CPS is present in all three conditions, as the same stimulus conditions were present prior to this IPh as were found prior to the first IPh in condition B, and a clear positive going shift occurs during the pause for all conditions.

Condition B, which contained an earlier IPh and pause, consistently contained shorter pauses during this TW both prior to and following editing for consistency. Pannekamp et al., (2005), from whom the stimuli in the present study were adapted, provided a chart showing that the pause in this condition was substantially shorter than that in the contrasting condition. The reason for this difference is unclear, but may be due to the higher potential for contrasting interpretations at the first IPh. The duration of the pause prior to *and be a good boy for a while* does not signal the presence of syntactic ambiguity, contrary to the pause following the first IPh. Thus, the clarity provided by a long intonational break is less crucial for the correct interpretation of the sentence. The duration of this pause (200 ms) led to most of the differences between conditions, as the following phrase began shortly after the CPS began to peak, causing a return to baseline. In response to conditions A and C a CPS is observed, but responses to condition C contrast with those to condition A, as a result of the compounding of P600 and CPS.

The comparisons between all three conditions show a clear CPS response in the ASD group during the second IPh (CPS2). When comparing conditions A and C in this TW, the primary differences found for the ASD group are maximal at anterior electrodes around 100 ms after the onset of the pause. This resulted in a significant difference between conditions as a result of the anterior/posterior distribution, as the anterior effect is only found in the ASD group. These differences can be interpreted as a negativity in condition A, a positivity in condition C, or both.

Recent studies (Koolen, et al., 2013; Koolen, et al., 2014) have described a Sustained Anterior Negativity (SAN) in ASD individuals in response to linguistic stimuli that contain semantic and orthographic violations, but, as with the P600, only when the ASD participants are instructed to attend to the plausibility of such sentences. In the present study, participants were instructed to rate sentences based on naturalness, which is a similar judgement to plausibility, but can contain other facets of linguistics and acoustics as well. The effect observed here, however, crucially differs, as there appears to be a negativity in condition A, the condition without the violation, while no such negativity is apparent in condition C. Further, the effect is not sustained, but would be more appropriately described as a negative peak. Thus, the negativity observed in this condition cannot be attributed to the

SAN.

It seems also unlikely that this negative peak is actually a delayed N400. This would mean that the effect occurs closer to 600 ms following the onset of the verb, *in the condition with no violation.* While previous studies have observed a delayed N400 in autistic participants, (Valdizán, et al., 2003), it seems highly unlikely that an N400 would only occur in the absence of any type of violation.

A more plausible scenario is that the early negative peak following the pause is a composite N1 offset response to the transition from acoustic stimulus to a gap in noise. Indeed, in the earlier CPS, a negative spike can also be seen in central and parietal electrodes peaking around 100 ms for the ASD group. A heightened sensitivity to the acoustic features of a silenced pause contrasting with vocal noise may cause this offset effect, and the less robust prosodic processes indexed by the CPS may allow its presence to be observed in this group, as opposed to the NT group where no negativity can be seen. This spike can also be seen for the ASD group in condition B, with a nearly identical distribution and peak latency. Thus, in nearly every situation where a CPS was expected, a sharp early offset negativity accompanies it for the ASD group. Condition C, however, lacks this negativity during the CPS2 TWs.

One possibility is that this lack of an offset component for the ASD group in condition C is the result of a positivity in condition C summating with the negativity to obscure its presence. This occurs in the condition with the prosodic violation. As we have seen that the ASD group rated condition C as less natural than either of the other two conditions, it is reasonable to assume that a Garden Path effect is present in the ASD group. While peak-to-peak analysis showed an effect of condition in the Garden Path TWs for the NT group, no such effect was shown (at electrodes Cz and Pz) for the ASD group. As the P600 and CPS2 summate for the NT group in this TW, the logical interpretation is that the lack of a composite N1 effect in condition C is due to its cancelation by a P600 effect that is simply earlier and more anterior than that found in the NT group. The first 100 ms following the onset of the pause corresponds to between 575 and 675 ms following the onset of the crucial verb creating the incongruity. It seems likely that the P600 is occurring during this period in the ASD group, which is earlier than the P600 peak seen in the NT group. While the P600 summates with the CPS in the NT group to create significantly different peak amplitudes, it summates earlier in the ASD group with the offset N1 effect. The reason for this earlier latency and more anterior distribution is unclear, and more research is needed to determine the exact cause and more accurately describe this effect. Regardless of the interpretation of these results, no Garden Path effect is seen for the ASD group at the expected posterior sites, likely due to its distribution being more anterior.

Further, the P600 that is observed in the ASD group has a much more left-lateralized distribution. This contrasts starkly with responses to IPh boundaries which contain no violation, as CPS responses are almost exclusively found to be right-lateralized in these contexts. The P600 effect shows a more left-hemisphere distributed (as well as more anterior) in contrast to the CPS responses, but both hemispheres are effected, suggesting that this response is fairly widely distributed across both hemispheres at anterior sites. The difference in distribution between conditions suggests that while responses to prosody on its own are limited to the right hemisphere, once the syntactic violation occurs, left-hemisphere syntactic processing areas become more involved in attempting to resolve the incongruity. Thus, the prosodically cued syntactic violation incorporates left-hemisphere areas. While prosody on its own is processed for this group almost solely in the right hemisphere, the incorporation of left hemisphere areas in response to a violation suggests that this group is indeed capable of reconciling information from the two separate pathways, but that they only make use of the consolidation strategy in an attempt to resolve prosodically driven syntactic incongruities.

4.2 Music

4.2.1 Behavioral Response

As was expected, no significant differences emerged in the behavioral responses to the music trials. The presence of a rest at a phrase-boundary is a common technique in musical phrasing, yet it is not uncommon for fills to be found during phrase-boundaries as well. Composers and musicians often insert embellishments during this period as a matter of taste and to add variation or personality to the musical form. Thus, a naturalness ranking for a phrase-boundary with different, yet still syntactically acceptable, fills would not be expected to differ greatly for any group. The task was chosen in an attempt to provide consistency across language and music domains, while maintaining participants' attention during the trials without being too demanding for those with no musical training.

4.2.2 EEG

Phrase Boundary Responses: Neuro-Typical Group

The present study is the first to identify a clear pattern of a sustained positive shift during the phrase boundary for both NT and ASD groups in response to the Phrased condition. This shift is defined by an anteriorly maximal spike, followed by a sustained positivity lasting for the duration of the rest. Once the first post-boundary note occurs, the onset components again line up together, obscuring any possible lasting effects of this positivity. This effect was observed using baselines from the last period during which all three conditions were identical, the 200 ms prior to the onset of the phrase-final note, (-2000 to -1800 relative to the post-boundary note onset), as well as a more proximate baseline of -1800 to -600 ms relative to the post-boundary note onset, however, only the more proximate baseline resulted in significant differences between responses for the NT group.

Phrase Boundary Responses: ASD Group

While the ASD group shows a similar response to the NT group to the Phrased condition, the ASD group does not show the same degree of difference between responses to conditions N and P in this TW. Observation of the grand average waveforms shows that a positive spike exists at the offset of the note in condition P, followed by a sustained positivity, similar to the response observed in the NT group. In the NoPause condition (N), a positive drift was also observed for the ASD group, unlike for the NT group. The presence of this drift suggests that phrase-final pre-boundary notes elicit positive potentials in a wider variety of phrase-boundary contexts, such as those which consist solely of a long held note. Together, these results show that a similar processing response occurs in the ASD group to both conditions.

If we interpret the presence of these responses as being representative of phrase-boundary processes, it suggests that the ASD group more accurately perceives the phrase boundary in the condition with sustained phrase-final notes than the NT group. Considering the common autistic affinity for music, it would be hardly surprising to observe similarities in neuro-physiological processing to a neuro-typical group. The presence of an additional processing ability, however, remains surprising. This suggests that not only is musical processing ability preserved in the ASD group, but that perhaps atypical neural organization provides this group with some additional benefit for musical processing, possibly due to reversed asymmetry of IFG.

This is not the first study to propose preserved or even heightened ASD musical ability. Lai, et al., (2012) showed enhanced LIFG activation in a low-functioning ASD group in response to music as opposed to language. The reasons for this are, however, unclear, but the presence of non-impaired phrase-boundary processing ability provides counter evidence to a general theory of deficient connectivity, or even local bias. Grouping regular metrical structures into coherent units is an ability that is relatively global, as it groups units of auditory input over the course of several seconds. It is important to note, however, that the distribution of this effect in both ASD and NT groups is similar to the distribution of the language CPS in the ASD group. Both responses were frontally distributed with less of an effect more posteriorly. The lateralization of the two effects was different, however, as the response to the musical phrase boundaries were distributed across hemispheres, while the language CPS was observed almost entirely in the right hemisphere. Further research must be undertaken to replicate and expound upon these preliminary results.

Origin of Phrase Boundary Effect

In order to examine the similarities between language and music phrase boundary processing mechanisms, it was necessary to create the proper acoustic and structural environment at these musical phrase-boundaries to examine the neurological processing responses. A boundary filled with notes results in electrophysiological responses containing a high incidence of auditory onset components obscuring higher order components with longer latency and smaller amplitude. Previous studies have only compared responses to musical stimuli with either a 'Phrased', or 'UnPhrased' (less phrased) designation. This comparison lacks an effective control, as the responses to the two conditions will differ greatly simply as a result of a greatly differing acoustic environment. Thus, seeking to examine components found during the phrase-boundary will be inevitably fraught with improper comparisons. For this reason, many previous studies (e.g. Neuhaus, et al., 2006; Nan, et al., 2009) have failed to identify any boundary related processing components *during* the phrase-boundary (but see Silva, et al., 2014). By creating an additional condition, the NoPause condition, we were able to compare responses to a condition containing a clear pause or rest (Phrased) to a condition with the same note sustained across the phrase-boundary. This allowed for a comparison of a true phrase-boundary with the commonly identified musical features of extended final note followed by a pause, with a condition in which the pause itself was lacking.

All conditions contain at least partial phrase-boundaries in the present study. There were no conditions that contained an absence of phrase boundary due to the presence of additional phrase signaling structural characteristics. For example, The phrase-final note was in all conditions the longest note in the musical piece, which similar to syllabic lengthening in language, helps to signify the end of a musical phrase. While in language, an intonational direction change is commonly found phrase-finally (Selkirk, et al., Steinhauer et al., 1999), examples of musical phrases ending with both upward and downward movement are common.

Larger structural features of musical phrasing were also present in the present stimuli. These include the metrical regularity of 4/4 time dictating predictable strong and weak beat stress assignment. (Lerdahl & Jackendoff, 1977; 1983) The phrase-final note always occurred on the first beat of a measure, which (at least in most Western musical systems) tends to be the strongest beat of a phrase. This beat aids in signaling the ends of phrases, particularly when other contextual features are present, such as a lengthened final note, and the presence of a subsequent rest. Additionally, this beat consistently occurred at the end of 4, 8, 12, or 16 complete bars. The division of musical pieces into even phrases with equal length is one of the features of most musics worldwide, and serves as another cue for phrase-boundary recognition. (Krumhansl & Jusczyk, 1990) These equal phrases provide structural regularity to the metrical hierarchy of musical systems. (Lerdahl & Jackendoff, 1983) Additionally, the phrase boundaries were approximately 10 seconds removed from each other, eliminating the likelihood for interference of phrase-boundaries occurring in the immediate temporal proximity of one another. Thus, the musical pieces served as effective presentations of musical phrase-boundaries at which to observe electrophysiological processing responses.

Unlike the language stimuli where the pauses were silenced in order to remove potential confounds of phrase-boundary processing such as breaths, the pauses in the music stimuli were not silenced. As described in the Methods section, there were highly significant differences between mean intensities during this period across all conditions, with Phrased being the quietest, UnPhrased being the loudest, and with NoPause falling around halfway in between. This comparison allowed for a relatively natural investigation of phrase-boundary processes, as if the observed responses were to occur only in conditions of absolute silence, their informative potential would be severely limited, and likely not representative of musical phrase boundary processing in natural contexts. The stimuli are thus fairly natural representations of music as heard in more typical circumstances, aside from the piano sound being a realistic synthesization. As a result, we can conclude that while acoustic features of pause recognition are present, such as the sharp spike observed in the Phrased condition at the offset of the final note, responses following this spike represent higher-level cognitive processes related to phrase-boundary processing.

While the distribution of this effect is more anterior than the language CPS, its time-course is more similar than the previously described ‘music CPS’, as it occurs during the phrase-boundary, as opposed to 500-600 ms following the first post-boundary note. Additionally, the fact that it is a sustained positive shift, as opposed to being a single temporally concise positivity, provides further support for this component being the musical equivalent to the language CPS. The question remains, however, what this effect signifies. The CPS in language is considered to be a chunking mechanism responsible for consolidating the structural and phonological features of phrases. This electrophysiological positivity in response to music likely represents similar chunking processes. The different distribution is likely due to the different, but overlapping cell networks involved in the processing of music and language. That the duration and latency remain almost identical suggests similar underlying operations responsible for the effect. Questions remain, however, regarding what contextual components are necessary to cue this response. Is it due to the processing of a structurally regular and predictable phrase-boundaries, or would any lengthened note followed by a rest be sufficient to elicit it? That it has yet to be described in the literature despite decades of auditory EEG/ERP research suggests that it is not a purely acoustic effect. The fact that a slower drift whose positivity lacks the sharp peak of the Phrased condition can be observed in the NoPause condition, particularly in the ASD group, suggests that this effect is indeed related to structural processing in a harmonic context.

Music CPS

In the present study, no 'music CPS' was observed for either group for any of the conditions, with any of the baselines used. Several studies in which the 'music CPS' has been found (Knösche, et al., 2005; Neuhaus, et al., 2006; Nan, et al., 2009) mention using the more distant baseline to confirm the results found from the use of a closer baseline, but none have included the data in the published papers corresponding to the studies, so we can only assume that these previous studies found consistent results by using multiple baseline comparisons. The present study, with the use of these multiple baselines consistently found no 'music CPS'. In fact, while there are statistical differences between conditions in this TW using the baseline of -200 to 0, they are almost all due to relative positivity in the UnPhrased conditions.

The previously described 'music CPS' may be simply due to auditory evoked potentials (AEPs: P2, or P200, in this case) of greater amplitude in the Phrased condition, as these components are consistently larger following periods of silence (longer ISI) than when following a period of other sounds. This is an effect of the refractory period of neural cell assembly response to the short ISI between repeated sounds. (Budd et al., 1998; Rosburg, et al., 2010) In the case of the 'music CPS', the positivity observed could simply be the result of a larger P200 following a period of silence in the Phrased condition during which the cell assembly comes closer to completing its refractory period. While Neuhaus et al., (2006) address this possibility, they only mention that it is unlikely that the CPS is actually the *first* P200 following the pause. That these previous studies did not have consistent note lengths in the stimuli prevents a reader from determining whether additional notes were present in this TW, and thus determining if this positivity is due to a sequential note. The authors mention that the distribution of the CPS is more posterior contrary to the anterior distribution of the P200, yet both are maximal at Cz.

The first comparison of musicians and non-musicians by Neuhaus et al. (2006) also found no 'music CPS' in non-musicians, attributing this lack of effect to the increased cognitive demand of the task. Since the same study found that P200 responses were much larger for musicians, it may be that this is the effect that was actually observed. While Nan, et al., (2009) found that non-musicians do actually elicit this effect when task requirements involve lighter cognitive load, the absence of this effect in the Neuhaus study suggests that the cognitive phenomena indexed by this component are quite specific. Several additional studies have also failed to find the 'music CPS' (e.g. Istók, et al., 2013; Hoshi-Shiba, et al., 2014) in expected environments, suggesting that its presence may be due to a potential confound, or that it is a phenomenon extremely sensitive to musical environment, and thus not broadly informative about cognitive musical phrase structure processing.

In the present study, the UnPhrased condition shows AEP effects with greatly reduced amplitude following the phrase-boundary in comparison to the other conditions. Despite this reduced amplitude, responses to this condition (U) tended to be more *positive* than the other conditions (N, P) in the TW of 500 to 600 ms, particularly at anterior electrodes, largely due to a large negative spike in the other conditions around 400 ms. A similar negativity was observed in non-musicians in Neuhaus et al., (2006), but in this case the likelihood of it being due to an overly attention demanding task is highly unlikely. The identity of this negativity is uncertain, although it may be simply a result of the baseline chosen.

There were many significant effects both in the global ANOVA and in the pairwise comparisons for this TW, however, when the baseline of -200 to 0 ms was used. The presence of a large positivity during the period of baseline correction for the Phrased condition meant that the comparisons were offset by an effect in one condition (P), and to a lesser extent in another (N), but absent in the third (U). Thus, any conclusions drawn from the use of this baseline comparison in the previously described ‘music CPS’ TW must be tentative at best. In any case, no positivity, or ‘music CPS’ was observed following the phrase-boundary for the Phrased or NoPause conditions.

Cadence and Repetition Effects

There were two TWs chosen based on visual inspection of the grand average waveforms that required closer examination. These were the TWs from -1600 to -1300, and -1200 to -600. In these comparisons there were significant differences between conditions N/P and condition U, but not between conditions N and P. These differences are due to a relative negativity for the ASD group during these TWs in condition U, and a positive drift in the other conditions. This suggests that the ASD group is responding to some difference during the phrase-final note of condition U. The notes themselves do not differ, which leaves the possibility that when listening to the second, third, and fourth phrases there is some expectation of a phrase-boundary filled with notes, with an expectation for the notes’ non-existence in the other conditions. As there are no significant repetition effects for this TW, this interpretation remains speculative, and the cause of these differences remains unknown. The visual inspection of these TWs confirms, however, that there is a continual anterior positive drift in all conditions as the phrase-final note is held. This positive drift is less marked in condition U, yet still exists, suggesting that phrase-final notes may elicit a positive drift in all conditions, that becomes more well defined as the note ends.

Regarding the effect of Cadence type on electrophysiological response, most significant results were found in the TW previously examined in ‘music CPS’ studies +450 to +600 ms following the onset of the post-boundary note. Previous studies have described an effect of cadence type (tonic vs. dominant) on the ‘music CPS’ in the limited situation of a proximate baseline, and therefore we sought to replicate or expand these results to gain insight into the effect of harmonic closure type on the CPS. In the present study, the comparison of the two cadence types (half vs. full) was confounded by the presence of a two-note anacrusis (two eighth notes) in some (3) of the stimuli, while other stimuli contained only one quarter note (9). Additionally, the final cadences at the end of the piece were always followed by a single quarter note on the first scale degree, regardless of the number of notes present in the earlier anacruses in the musical piece. In other words, the comparison of post-boundary notes in the full cadence condition was unevenly split between one- and two-note environments. While for examining the presence of the music CPS, this feature of the stimuli was balanced across conditions N v P v U, that it is only imbalanced in the full cadential phrase-boundaries unfortunately limits the potential for meaningful conclusions interpreted from the analysis of this TW.

In terms of the effects of repetition on the phrase-boundary effect as examined by comparing responses to the first and second presentations of a musical phrase due to the repeat in the musical pieces, different results for the different groups resulting in significant Cond x Grp interactions require different interpretations. For the NT group, the decrease in phrase-boundary response to repeated presentations, particularly at posterior and lateral electrodes, suggests efficiency of neural resource allocation. On repeated presentations, less attention and fewer neural resources need to be dedicated to the consolidation of musical phrases. The fact that this group has already processed the musical structure of the piece following the first presentation means that cognitive resources, including attention and memory operations, can be reallocated to some other purpose. Any incongruous features of the stimulus will be then perceived through mismatch, or oddball, detection and exogenous orienting. Further, the need to make predictions about what will come in the future is diminished on repeated presentation of the same stimulus, particularly since following the presentation of the first few musical pieces, participants likely expect that all the pieces will generally follow the same structure.

For the ASD group, on the other hand, increases in phrase-boundary effect size were observed at both posterior and lateral electrodes in response to the repeated presentations of musical pieces. Considering that repetitive behavior is one of the diagnostic criteria of autism, there may be some comfort found in repetition for individuals with this type of disorder. Perhaps in the case of music, repeated presentation allows individuals with this disorder to focus less on the local details of individual notes and tones, to more global features of structure and phrasing. Consolidation of global information is far from impossible for individuals with these disorders, thus, the cognitive model is known as local *bias*. Perhaps this bias is especially strong on initial presentations, while following repeated presentations, global details about a stimulus gradually become apparent. While for neuro-typicals, repeated presentation of a stimulus no longer requires the same level of attention and processing of structural features (as in the present study) becomes more automatic. Perhaps in ASD individuals, repeated presentation allows local features of the stimulus to be processed more automatically, freeing up cognitive resources for global consolidation. If this were the case, then it would help explain the common ASD preference for familiar routines, music, and perhaps even behaviors. It may also explain the prosodic impairment for language.

The fact that one of the central features of music is that melodies, rhythms, and structures repeat may provide a clue as to why music and prosody, despite generally sharing similar acoustic features and neural processing resources, show such a marked dissociation in ASDs. The prosody of an utterance is rarely reproduced exactly, while musical pieces, particularly recorded music, frequently repeat phrases with a high-degree of exactness in terms of pitch, note order, rhythm, tempo, and pattern of notes. Perhaps the high degree of local information in language prevents global prosody from being incorporated as a meaningful element of the speech stream on first listen. With the exception of recorded speech, no second listen is generally available, leaving ASD individuals without a chance to more automatically attend to local features and focus on global prosody.

5 Conclusions

The present study is the first to study the electrophysiological prosodic processing responses of individuals with autism spectrum disorders (ASDs). The significant between-group differences in response to the language conditions confirms previous behavioral results demonstrating a prosodic processing impairment in this group. Additionally, the present study provides electrophysiological evidence to support these results, which shows not only decreased response to prosodic breaks, but changes in the distribution of responses, suggesting that this impairment is due to altered neurophysiological organization.

The reason for this impairment appears to be the fact that attending to phrase level cues designed to aid in parsing prosodic boundaries is impaired by a heightened sensitivity to low-level, or local, acoustic features of auditorily presented language stimuli. These conclusions support the theories of functional under-connectivity, weak central coherence, and local bias, as an overabundance of short range connections at the expense of long range frontal-posterior or inter-hemispheric connections would predict the interference of local auditory processing response on more global prosodic mechanisms. The more anterior distribution for the ASD group of what is generally a posteriorly maximal response provides additional support to frontal-posterior under-connectivity. The heterogeneity of the disorder, however, creates variability within the participant pool, and all conclusions regarding this group are statistical tendencies, as opposed to strict, definitive universals.

This is also the first study to show a clear positive shift during the musical phrase boundary similar to the prosodic processing component, the Closure Positive Shift found in language, in both neuro-typical and ASD groups. This positivity likely serves the same function as the CPS observed in response to linguistic/prosodic phrases. It appears to represent a chunking mechanism that cues hierarchical and metrical structuring of the musical input into cohesive phrases, allowing the listener to organize the previously heard musical information while making predictions about what is to follow. This effect supports the well-established theories regarding shared music and language neuro-physiological processing resources. The presence of similar responses from the ASD group to two conditions not only supports the preserved musical ability in this disorder, but suggests that individuals with this disorder may actually demonstrate heightened musical processing ability. While previous studies have found similar results, more research is needed to precisely define and more thoroughly describe what these abilities may entail.

The absence of a 'music CPS' following the onset of the post-boundary phrase in the present study suggests that this effect may be extremely sensitive to musical context, or caused by a confounding effect of phrase-boundary type. Multiple published studies (and maybe more unpublished ones) have failed to observe this effect in predicted environments. Together with the facts that neither the latency, duration, nor the morphology of the component, despite being a posterior positivity, resemble the CPS observed in language suggests that this effect is driven by some other cognitive process, perhaps relating to the onset of phrases.

The question remains: why does a dissociation exist between prosodic and musical processing abilities in ASDs? Both abilities rely on the complex integration of wide networks of cell assemblies incorporating various left and right hemisphere structures responsible for acoustic perception and rule-based operations. One explanation is that hemispheric asymmetry of inferior frontal areas allows for preserved or heightened musical processing abilities due to an over-abundance of short range connections between areas responsible for auditory perception and areas responsible for broader integration of that input. Consolidating prosodic information in language, however, involves precise and efficient coordination of the dual language pathways in both hemispheres. A deficit of long-range inter-hemispheric connections may result in an impairment consolidating global prosodic information from the right hemisphere with global lexical/semantic representations and syntactic operations in the left hemisphere. While more research is required to explicitly test predictions made from under-connectivity theory, the present study has used it merely as a descriptive framework in order to more thoroughly understand the linguistic deficit in autism. Many of the conclusions from the present study support this theory, however, it does not account for preserved musical ability in autism.

Another explanation of the dissociation between music and language in autism is that the repetition found in music provides ASD individuals the chance to automatically process low-level, local information, while freeing up higher cognitive resources to attend to global features. The lack of exact repetition in prosody may prevent the global features of prosody from ever being processed. Further research is needed, however, to specify the neurological prosodic consolidation deficit characteristic of ASDs, the preserved, or even heightened nature of autistic music processing ability, and to explore the nature of repetition in ASD cognition. Further research is also needed to address therapeutic avenues to reduce this prosodic impairment in order to improve social and linguistic functioning for these individuals.

6 References

Aasland, W. A. and Baum, S. R. (2003). Temporal parameters as cues to phrasal boundaries: A comparison of processing by left- and right-hemisphere brain-damaged individuals. *Brain and Language*, 87(3):385–399.

Accordino, R., Comer, R., and Heller, W. B. (2007). Searching for music's potential: A critical examination of research on music therapy with individuals with autism. *Research in Autism Spectrum Disorders*, 1(1):101–115.

Amaral, D. G. (2011). The promise and the pitfalls of autism research: an introductory note for new autism researchers. *Brain Research*, 1380:3–9.

Ames, C. and Fletcher-Watson, S. (2010). A review of methods in the study of attention in autism. *Developmental Review*, 30(1):52–73.

APA (2000). *Diagnostic & Statistical Manual of Mental Disorders: DSM-IV-TR*. Washington, DC, 4th edition.

Asperger, H. (1944). Die „Autistischen Psychopathen“ im Kindesalter. *European Archives of Psychiatry and Clinical Psychology*, pages 1–64.

Astésano, C., Besson, M., and Alter, K. (2004). Brain potentials during semantic and prosodic processing in French. *Brain Research*, 18(2):172–184.

Baron-Cohen, S. (1995). *Mindblindness: An Essay on Autism and Theory of Mind*. MIT Press, Cambridge, Mass.

Baron-Cohen, S., Leslie, A., and Frith, U. (1985). Does the autistic child have a "theory of mind"?*. *Cognition*, 21:37–46.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., and Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, 31(1):5–17.

Beckman, M. E. and Venditti, J. J. (2011). Intonation. In Goldsmith, J., Riggle, J., and Yu, A. C., editors, *The Handbook of Phonological Theory*, chapter 15, pages 495–532. Blackwell Publishing Ltd., 2nd edition.

Boersma, P. and Weenink, D. (2015). Praat: Doing phonetics by computer (version 5.3.39).

Bonnel, A., Mottron, L., Peretz, I., Trudel, M., Gallun, E., and Bonnel, A.-M. (2003). Enhanced pitch sensitivity in individuals with autism: a signal detection analysis. *Journal of Cognitive Neuroscience*, 15(2):226–35.

Buday, E. M. (1995). The effects of signed and spoken words taught with music on sign and speech imitation by children with autism. *Journal of Music Therapy*, 32:189–202.

Budd, T., Barry, R., and Gordon, E. (1998). Decrement of the N1 auditory event-related potential with stimulus repetition: habituation vs. refractoriness. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, 31:51–68.

Caria, A., Venuti, P., and de Falco, S. (2011). Functional and dysfunctional brain circuits underlying emotional processing of music in autism spectrum disorders. *Cerebral Cortex*, 21(12):2838–49.

Chevallier, C., Noveck, I., Happé, F., and Wilson, D. (2011). What's in a voice? Prosody as a test case for the Theory of Mind account of autism. *Neuropsychologia*, 49(3):507–17.

Corbetta, M. and Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews: Neuroscience*, 3(3):201–15.

Corbetta, M., Shulman, G. L., Miezin, F. M., and Petersen, S. E. (1995). Superior parietal cortex activation during spatial attention shifts and visual feature conjunction. *Science*, 270(5237):802–5.

Cowles, H. W., Kluender, R., Kutas, M., and Polinsky, M. (2007). Violations of information structure: an electrophysiological study of answers to wh-questions. *Brain and Language*, 102(3):228–42.

de Angulo, J. (1929). Grammatical Processes: Incremental vs. Autonomic. *Language*, 5(2):117–118.

De Fossé, L., Hodge, S. M., Makris, N., Kennedy, D. N., Caviness, V. S., McGrath, L., Steele, S., Ziegler, D. a., Herbert, M. R., Frazier, J. a., Tager-Flusberg, H., and Harris, G. J. (2004). Language-association cortex asymmetry in autism and specific language impairment. *Annals of neurology*, 56(6):757–66.

Eckstein, K. and Friederici, A. D. (2006). It's early: event-related potential evidence for initial interaction of syntax and prosody in speech comprehension. *Journal of Cognitive Neuroscience*, 18(10):1696–711.

Edgerton, C. (1994). The effect of improvisational music therapy on the communicative behaviors of autistic children. *Journal of Music Therapy*, 31:31–62.

Floris, D. L., Chura, L. R., Holt, R. J., Suckling, J., Bullmore, E. T., Baron-Cohen, S., and Spencer, M. D. (2013). Psychological correlates of handedness and corpus callosum asymmetry in autism: the left hemisphere dysfunction theory revisited. *Journal of Autism and Developmental Disorders*, 43(8):1758–72.

Fodor, J. (2000). *The mind doesn't work that way: The scope and limits of computational psychology.* MIT Press, Cambridge, Mass.

Foerstl, J. (1989). Early interest in the idiot savant. *The American Journal of Psychiatry*, 146(4).

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, 6(2):78–84.

Friederici, A. D. and Alter, K. (2004). Lateralization of auditory language functions: a dynamic dual pathway model. *Brain and Language*, 89(2):267–76.

Frith, C. (2003). What do imaging studies tell us abut the neural basis of autism? In Bock, G. and Goode, J., editors, *Autism: Neural Basis and Treatment Possibilities: Novartis Foundation Symposium*, volume 251. John Wiley & Sons.

Frith, U. (1989). *Autism: Explaining the Enigma.* Basil Blackwell, Oxford, UK.

Frith, U. and Happé, F. (1994). Autism: beyond "theory of mind". *Cognition*, 50(1-3):115–32.

Garza Villarreal, E. A., Brattico, E., Leino, S., Ostergaard, L., and Vuust, P. (2011). Distinct neural responses to chord violations: a multiple source analysis study. *Brain Research*, 1389:103–14.

Golan, O., Baron-Cohen, S., Hill, J. J., and Rutherford, M. D. (2007). The 'Reading the Mind in the Voice' test-revised: a study of complex emotion recognition in adults with and without autism spectrum conditions. *Journal of Autism and Developmental Disorders*, 37(6):1096–106.

Gordon, R. L., Magne, C. L., and Large, E. W. (2011). EEG Correlates of Song Prosody: A New Look at the Relationship between Linguistic and Musical Rhythm. *Frontiers in Psychology*, 2(November):352.

Grandin, T. (1996). *Thinking in Pictures.* Vintage, New York.

Hallmayer, J., Cleveland, S., Torres, A., Phillips, J., Cohen, B., Torigoe, T., Miller, J., Fedele, A., Collins, J., Smith, K., Lotspeich, L., Croen, L. A., Ozonoff, S., Lajonchere, C., Grether, J. K., and Risch, N. (2011). Genetic heritability and shared environmental factors among twin pairs with autism. *Archives of General Psychiatry*, 68(11):1095–102.

Happé, F. and Frith, U. (2006). The weak coherence account: detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 36(1):5–25.

Herbert, M. R., Harris, G. J., Adrien, K. T., Ziegler, D. A., Makris, N., Kennedy, D. N., Lange, N. T., Chabris, C. F., Bakardjiev, A., Hodgson, J., Takeoka, M., Tager-Flusberg, H., and Caviness, V. S. (2002). Abnormal asymmetry in language association cortex in autism. *Annals of Neurology*, 52(5):588–96.

Hesling, I., Dilharreguy, B., Peppé, S., Amirault, M., Bouvard, M., and Allard, M. (2010). The integration of prosodic speech in high functioning autism: a preliminary FMRI study. *PloS One*, 5(7):e11571.

Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience*, 21(7):1229–43.

Hickok, G. and Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*, 92(1-2):67–99.

Hodge, S. M., Makris, N., Kennedy, D. N., Caviness, V. S., Howard, J., McGrath, L., Steele, S., Frazier, J. A., Tager-Flusberg, H., and Harris, G. J. (2010). Cerebellum, language, and cognition in autism and specific language impairment. *Journal of Autism and Developmental Disorders*, 40(3):300–16.

Hoshi-Shiba, R., Furukawa, K., and Okanoya, K. (2014). Neural correlates of expectation of musical termination structure or cadence. *Neuroreport*, 25(10):743–8.

Janata, P. (1995). ERP measures assay the degree of expectancy violation of harmonic contexts in music. *Journal of Cognitive Neuroscience*, 7(2):153–164.

Jeste, S. S. (2011). The neurology of autism spectrum disorders. *Current Opinion in Neurology*, 24(2):132–9.

Just, M. A., Cherkassky, V. L., Keller, T. a., and Minshew, N. J. (2004). Cortical activation and synchronization during sentence comprehension in high-functioning autism: evidence of underconnectivity. *Brain: A Journal of Neurology*, 127(Pt 8):1811–21.

Just, M. A., Keller, T. A., Malave, V. L., Kana, R. K., and Varma, S. (2012). Autism as a neural systems disorder: a theory of frontal-posterior underconnectivity. *Neuroscience and Biobehavioral Reviews*, 36(4):1292–313.

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, 2:217–250.

Kerkhofs, R., Vonk, W., Schriefers, H., and Chwilla, D. J. (2007). Discourse, syntax, and prosody: the brain reveals an immediate interaction. *Journal of Cognitive Neuroscience*, 19(9):1421–34.

Knösche, T. R., Neuhaus, C., Haueisen, J., Alter, K., Maess, B., Witte, O. W., and Friederici, A. D. (2005). Perception of phrase structure in music. *Human Brain Mapping*, 24(4):259–73.

Koelsch, S. (2012). *Brain & Music*. John Wiley & Sons, Ltd., West Sussex.

Koelsch, S., Gunter, T., Friederici, A. D., and Schröger, E. (2000). Brain indices of music processing: "nonmusicians" are musical. *Journal of Cognitive Neuroscience*, 12(3):520–41.

Koelsch, S., Gunter, T., Schröger, E., and Friederici, A. D. (2003). Processing tonal modulations: an ERP study. *Journal of Cognitive Neuroscience*, 15(8):1149–59.

Koelsch, S., Gunter, T. C., Wittfoth, M., and Sammler, D. (2005). Interaction between syntax processing in language and in music: an ERP Study. *Journal of Cognitive Neuroscience*, 17(10):1565–77.

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., and Friederici, A. D. (2004). Music, language and meaning: brain signatures of semantic processing. *Nature Neuroscience*, 7(3):302–7.

Koelsch, S., Rohrmeier, M., Torrecuso, R., and Jentschke, S. (2013). Processing of hierarchical syntactic structure in music. *Proceedings of the National Academy of Sciences of the United States of America*, 110(38):15443–8.

Koelsch, S. and Siebel, W. a. (2005). Towards a neural basis of music perception. *Trends in Cognitive Sciences*, 9(12):578–84.

Koolen, S., Vissers, C. T. W. M., Egger, J. I. M., and Verhoeven, L. (2013). Can monitoring in language comprehension in Autism Spectrum Disorder be modulated? Evidence from event-related potentials. *Biological Psychology*, 94(2):354–68.

Koolen, S., Vissers, C. T. W. M., Egger, J. I. M., and Verhoeven, L. (2014). Monitoring in language perception in high-functioning adults with autism spectrum disorder: evidence from event-related potentials. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 125(1):108–23.

Kotz, S. A., Meyer, M., and Paulmann, S. (2006). Lateralization of emotional prosody in the brain: an overview and synopsis on the impact of study design. *Brain Research*, 156(1992):285–94.

Kotz, S. A. and Paulmann, S. (2007). When emotional prosody and semantics dance cheek to cheek: ERP evidence. *Brain Research*, 1151:107–18.

Krumhansl, C. L. and Jusczyk, P. W. (1990). Infants' perception of phrase structure in music. *Psychological Science*, 1(1):70–73.

Kutas, M. and Hillyard, S. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207(4427):203–205.

Ladefoged, P. (2006). *A Course in Phonetics.* Harcourt Brace Janovich, New York, 5th. edition.

Lai, G., Pantazatos, S. P., Schneider, H., and Hirsch, J. (2012). Neural systems for speech and song in autism. *Brain: A Journal of Neurology*, 135(Pt 3):961–75.

Le Bel, R., Pineda, J., and Sharma, A. (2009). Motor-auditory-visual integration: The role of the human mirror neuron system in communication and communication disorders. *Journal of Communication Disorders*, 42(4):299–304.

Lehrl, S., Triebig, G., and Fischer, B. (1995). Multiple choice vocabulary test MWT as a valid and short test to estimate premorbid intelligence. *Acta Neurologica Scandinavica*, 91(5):335–345.

Lepistö, T., Kajander, M., Vanhala, R., Alku, P., Huotilainen, M., Näätänen, R., and Kujala, T. (2008). The perception of invariant speech features in children with autism. *Biological Psychology*, 77(1):25–31.

Lepistö, T., Kujala, T., Vanhala, R., Alku, P., Huotilainen, M., and Näätänen, R. (2005). The discrimination of and orienting to speech and non-speech sounds in children with autism. *Brain Research*, 1066(1-2):147–57.

Lerdahl, F. and Jackendoff, R. (1977). Toward a Formal Theory of Tonal Music. *Journal of Music Theory*, 21(1):111–171.

Lerdahl, F. and Jackendoff, R. (1983). *A Generative Theory of Tonal Music.* MIT Press, Cambridge, Mass.

Levitin, D. J. (2008). *The world in six songs: How the musical brain created human nature.* Dutton, New York.

Liss, M., Mailloux, J., and Erchull, M. J. (2008). The relationships between sensory processing sensitivity, alexithymia, autism, depression, and anxiety. *Personality and Individual Differences*, 45(3):255–259.

Liss, M., Saulnier, C., Fein, D., and Kinsbourne, M. (2006). Sensory and attention abnormalities in autistic spectrum disorders. *Autism: The International Journal of Research and Practice*, 10(2):155–72.

Lord, C., Risi, S., and Lambrecht, L. (2000). The Autism Diagnostic Observation Schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, 30(3):205–233.

Lord, C., Rutter, M., and Couteur, A. (1994). Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, 24:659–685.

Magne, C., Astésano, C., Lacheret-Dujour, A., Morel, M., Alter, K., and Besson, M. (2005). On-line processing of "pop-out" words in spoken French dialogues. *Journal of Cognitive Neuroscience*, 17(5):740–56.

Maljaars, J., Noens, I., Jansen, R., Scholte, E., and van Berckelaer-Onnes, I. (2011). Intentional communication in nonverbal and verbal low-functioning children with autism. *Journal of Communication Disorders*, 44(6):601–14.

Männel, C. and Friederici, A. D. (2009). Pauses and intonational phrasing: ERP studies in 5-month-old German infants and adults. *Journal of Cognitive Neuroscience*, 21(10):1988–2006.

Männel, C. and Friederici, A. D. (2011). Intonational phrase structure processing at different stages of syntax acquisition: ERP studies in 2-, 3-, and 6-year-old children. *Developmental Science*, 14(4):786–98.

Mithen, S. (2006). *The singing neanderthals: The origins of music, language, mind, and body.* Harvard University Press, Cambridge, Mass.

Mundy, P. and Gomes, A. (1998). Individual differences in joint attention skill development in the second year. *Infant Behavior and Development*, 21(3):469–482.

Nan, Y., Knösche, T. R., and Friederici, A. D. (2009). Non-musicians' perception of phrase boundaries in music: A cross-cultural ERP study. *Biological Psychology*, 82(1):70–81.

Nan, Y., Knösche, T. R., Zysset, S., and Friederici, A. D. (2008). Cross-cultural music phrase processing: an fMRI study. *Human Brain Mapping*, 29(3):312–28.

Norton, A., Zipse, L., Marchina, S., and Schlaug, G. (2009). Melodic intonation therapy: shared insights on how it is done and why it might help. *Annals of the New York Academy of Sciences*, 1169:431–6.

Oberman, L. and Ramachandran, V. (2007). The simulating social mind: The rold of the mirror neuron system and simulation in the social and communicative deficit of autism spectrum disorders. *Psychological Bulletin*, 133(2):310–327.

O'Connor, N. & Hermelin, B. (1984). Idiot savant calendrical calculations: math or memory? *Psychological Medicine*, 14:801–806.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The edinburgh inventory. *Neuropsychologia*, 9:97–113.

Organization, W. H. (1992). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines.* World Health Organization, Geneva.

Pannekamp, A. (2005). *Prosodische Informationsverarbeitung bei normalsprachlichen und deviantem Satzmaterial: Untersuchungen mit ereigniskorrelierten Hirnpotentialen.* MPI Series in Human Cognitive and Brain Sciences: 63.

Pannekamp, A., Weber, C., and Friederici, A. D. (2006). Prosodic processing at the sentence level in infants. *Neuroreport*, 17(6):675–8.

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, 6(7):674–81.

Patel, A. D. (2005). The relationship of music to the melody of speech and to syntactic processing disorders in aphasia. *Annals of the New York Academy of Sciences*, 1060:59–70.

Paul, R., Augustyn, A., Klin, A., and Volkmar, F. R. (2005). Perception and Production of Prosody by Speakers with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, 35(2):205–220.

Paulmann, S., Jessen, S., and Kotz, S. A. (2012). It's special the way you say it: an ERP investigation on the temporal dynamics of two types of prosody. *Neuropsychologia*, 50(7):1609–20.

Paulmann, S. and Kotz, S. A. (2008). An ERP investigation on the temporal dynamics of emotional prosody and emotional semantics in pseudo- and lexical-sentence context. *Brain and Language*, 105(1):59–69.

Paulmann, S., Pell, M. D., and Kotz, S. A. (2008). Functional contributions of the basal ganglia to emotional prosody: evidence from ERPs. *Brain Research*, 1217:171–8.

Peppé, S. and McCann, J. (2003). Assessing intonation and prosody in children with atypical language development: the PEPS-C test and the revised version. *Clinical Linguistics & Phonetics*, 17(4-5):345–354.

Peppé, S., McCann, J., and Gibbon, F. (2007). Receptive and Expressive Prosodic Ability in Children with High-Functioning Autism. *Journal of Speech, Language, and Hearing Research*, 50(August):1015–1028.

Peretz, I. (2006). The nature of music from a biological perspective. *Cognition*, 100(1):1–32.

Perkins, T., Stokes, M., McGillivray, J., and Bittar, R. (2010). Mirror neuron dysfunction in autism spectrum disorders. *Journal of Clinical Neuroscience: Official Journal of the Neurosurgical Society of Australasia*, 17(10):1239–43.

Pinker, S. (1997). *How the mind works.* Norton, New York.

Poulin-Charronnat, B., Bigand, E., Madurell, F., and Peereman, R. (2005). Musical structure modulates semantic priming in vocal music. *Cognition*, 94(3):B67–78.

Pratt, C. C. (1952). *Music as the language of emotion; a lecture delivered in the Whittall Pavilion of the Library of Congress, December 21, 1950.* U.S. Govt. Print Office, Washington, DC.

Riemann, H. (1900). *Musik-Lexikon.* M. Hesse, Leipzig, 5. vollstándig umgearb. aufl. ed. edition.

Rimland, B. (1964). *Infantile autism; the syndrom and its implications for a neural theory of behavior.* Appleton-Century-Crofts, New York.

Rimland, B. (1978). Savant capabilities of autistic children and their cognitive implications. In Serban, G., editor, *Cognitive Defects in the Development of Mental Illness.* Brunner/Mazel, New York.

Roach, P. (2000). *English Phonetics and Phonology.* Cambridge University Press, Cambridge, England.

Rohrmeier, M. (2011). Towards a generative syntax of tonal harmony. *Journal of Mathematics and Music*, 5(1):35–53.

Roll, M. and Horne, M. (2011). Interaction of right- and left-edge prosodic boundaries in syntactic parsing. *Brain Research*, 1402:93–100.

Rosburg, T., Zimmerer, K., and Huonker, R. (2010). Short-term habituation of auditory evoked potential and neuromagnetic field components in dependence of the interstimulus interval. *Brain Research*, 205(4):559–70.

Russo, N., Mottron, L., Burack, J. A., and Jemel, B. (2012). Parameters of semantic multisensory integration depend on timing and modality order among people on the autism spectrum: evidence from event-related potentials. *Neuropsychologia*, 50(9):2131–41.

Russo, N. M., Skoe, E., Trommer, B., Nicol, T., Zecker, S., Bradlow, A., and Kraus, N. (2008). Deficient brainstem encoding of pitch in children with Autism Spectrum Disorders. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 119(8):1720–31.

Rutherford, M. D., Baron-Cohen, S., and Wheelwright, S. (2002). Reading the mind in the voice: a study with normal adults and adults with Asperger syndrome and high functioning autism. *Journal of Autism and Developmental Disorders*, 32(3):189–94.

Sacks, O. and Olennick, N. (1995). *An anthropologist on mars*. Alfred A. Knopf, New York.

Sammler, D., Koelsch, S., Ball, T., Brandt, A., Grigutsch, M., Huppertz, H.-J., Knösche, T. R., Wellmer, J., Widman, G., Elger, C. E., Friederici, A. D., and Schulze-Bonhage, A. (2012). Co-localizing linguistic and musical syntax with intracranial EEG. *NeuroImage*, 64C:134–146.

Sammler, D., Kotz, S. A., Eckstein, K., Ott, D. V. M., and Friederici, A. D. (2010). Prosody meets syntax: the role of the corpus callosum. *Brain: A Journal of Neurology*, 133(9):2643–55.

Samson, F., Hyde, K. L., Bertone, A., Soulières, I., Mendrek, A., Ahad, P., Mottron, L., and Zeffiro, T. a. (2011). Atypical processing of auditory temporal complexity in autistics. *Neuropsychologia*, 49(3):546–55.

Schendel, D., Grø nborg, T., and Parner, E. (2014). The genetic and environmental contributions to autism: looking beyond twins. *JAMA: The Journal of the American Medical Association*, 311(17):1738–1739.

Schirmer, A. and Kotz, S. A. (2006). Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences*, 10(1):24–30.

Schirmer, A., Kotz, S. A., and Friederici, A. D. (2002). Sex differentiates the role of emotional prosody during word processing. *Brain Research*, 14(2):228–33.

Schmidt-Kassow, M. and Kotz, S. A. (2009). Event-related brain potentials suggest a late interaction of meter and syntax in the P600. *Journal of Cognitive Neuroscience*, 21(9):1693–708.

Selkirk, E. O. (1984). *Phonology and Syntax: The relation between sound and structure*. MIT Press, Cambridge, Mass.

Sigman, M. (1998). Joint Attention as a predictor of language skills and peer engagement in children with autism. *Infant Behavior and Development*, 21:144.

Silva, S., Branco, P., Barbosa, F., Marques-Teixeira, J., Petersson, K. M., and Castro, S. a. L. (2014). Musical phrase boundaries, wrap-up and the closure positive shift. *Brain Research*, 1585:99–107.

Soper, H. V., Satz, P., Orsini, D. L., Henry, R. R., Zvi, J. C., and Schulman, M. (1986). Handedness patterns in autism suggest subtypes. *Journal of Autism and Developmental Disorders*, 16(2):155–167.

Steinhauer, K. (2003). Electrophysiological correlates of prosody and punctuation. *Brain and Language*, 86(1):142–164.

Steinhauer, K., Alter, K., and Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience*, 2(2):191–6.

Steinhauer, K. and Friederici, A. D. (2001). Prosodic boundaries, comma rules, and brain responses: the closure positive shift in ERPs as a universal marker for prosodic phrasing in listeners and readers. *Journal of Psycholinguistic Research*, 30(3):267–95.

Strelnikov, K. N., Vorobyev, V. A., Chernigovskaya, T. V., and Medvedev, S. V. (2006). Prosodic clues to syntactic processing–a PET and ERP study. *NeuroImage*, 29(4):1127–34.

Toepel, U., Pannekamp, A., and Alter, K. (2007). Catching the news: Processing strategies in listening to dialogs as measured by ERPs. *Behavioral and Brain Functions: BBF*, 3:53.

Treffert, D. (1988). The idiot savant: a review of the syndrome. *American Journal of Psychiatry*, 145(5):563–572.

Valdizán, J., Abril-Villalba, B., M., M.-G., Sans-Capdevila, O., Pablo, M., Peralta, P., Lasierra, Y., and Bernal-Lafluente, M. (2003). Cognitive evoked potentials in autistic children. *Revista de neurologia*, 36(5):425–8.

Walenski, M., Mostofsky, S. H., and Ullman, M. T. (2014). Inflectional morphology in high-functioning autism: Evidence for speeded grammatical processing. *Research in Autism Spectrum Disorders*, 8(11):1607–1621.

Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C., and Winter, A. L. (1964). Contingent negative variation: an electric sign of sensorimotor association and expectancy in the human brain. *Nature*, 203(4943):380–384.

Wan, C. Y., Demaine, K., Zipse, L., Norton, A., and Schlaug, G. (2010). From music making to speaking: engaging the mirror neuron system in autism. *Brain Research*, 82(3-4):161–8.

Wenzlaff, M. and Clahsen, H. (2005). Finiteness and verb-second in German agrammatism. *Brain and Language*, 92(1):33–44.

Williams, J. H., Whiten, A., Suddendorf, T., and Perrett, D. I. (2001). Imitation, mirror neurons and autism. *Neuroscience and Biobehavioral Reviews*, 25(4):287–95.

Zatorre, R. J., Belin, P., and Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends in Cognitive Sciences*, 6(1):37–46.

Zellin, M., Pannekamp, A., Toepel, U., and van der Meer, E. (2011). In the eye of the listener: pupil dilation elucidates discourse processing. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, 81(3):133–41.

7 Appendices

7.1 Appendix I: Language Stimuli (Adapted from Pannekamp, et al., 2005)

01. Kevin verspricht Sophie zu schlafen und ganz lange lieb zu sein.
02. Kevin verspricht, Sophie zu küssen und ganz lange lieb zu sein.

03. Lena verspricht Sophie zu flitzen und Getränke zu kaufen.
04. Lena verspricht, Sophie zu helfen und Getränke zu kaufen.

05. Lena bittet Lukas zu tuten und lange rumzualbern.
06. Lena bittet, Lukas zu haschen und lange rumzualbern.

07. Sophie bittet Kevin zu kommen und das Spiel mitzumachen.
08. Sophie bittet, Kevin zu holen und das Spiel mitzumachen.

09. Lukas erlaubt Lena zu wüten und alles zu verraten.
10. Lukas erlaubt, Lena zu hauen und alles zu verraten.

11. Simba erlaubt Mietzi zu fauchen und die Vögel zu scheuchen.
12. Simba erlaubt, Mietzi zu baden und die Vögel zu scheuchen.

13. Lukas verbietet Kevin zu zappeln und vom Tisch aufzustehen.
14. Lukas verbietet, Kevin zu hänseln und vom Tisch aufzustehen.

15. Sophie verbietet Lukas zu nörgeln und ganz laut rumzuhupen.
16. Sophie verbietet, Lukas zu stören und ganz laut rumzuhupen.

17. Simba hilft Mietzi zu raufen und die Bälle zu schubsen.
18. Simba hilft, Mietzi zu ärgern und die Bälle zu schubsen.

19. Kevin hilft Lena zu lärmen und ganz laut zu klingeln.
20. Kevin hilft, Lena zu wecken und ganz laut zu klingeln.

21. Mietzi lehrt Simba zu rennen und die Katzen zu jagen.
22. Mietzi lehrt, Simba zu finden und die Katzen zu jagen.

23. Mietzi lehrt Simba zu schleichen und dabei aufzupassen.
24. Mietzi lehrt, Simba zu füttern und dabei aufzupassen.

25. Kevin verspricht Elke zu sitzen und am Tisch zu bleiben.
26. Kevin verspricht, Elke zu folgen und am Tisch zu bleiben.

27. Lena verspricht Elke zu saugen und das Geschirr zu spülen.
28. Lena verspricht, Elke zu stützen und das Geschirr zu spülen.

29. Elke bittet Kevin zu sausen und ganz schnell einzukaufen.
30. Elke bittet, Kevin zu suchen und ganz schnell einzukaufen.

31. Elke bittet Karsten zu tanken und eilig loszufahren.
32. Elke bittet, Karsten zu rufen und eilig loszufahren.

33. Karsten erlaubt Lena zu lachen und freudig loszukichern.
34. Karsten erlaubt, Lena zu kneifen und freudig loszukichern.

35. Mietzi erlaubt Simba zu schwimmen und den Kopf einzutauchen.
36. Mietzi erlaubt, Simba zu schnappen und den Kopf einzutauchen.

37. Simba verbietet Mietzi zu mauzen und die Krallen zu zeigen.
38. Simba verbietet, Mietzi zu strafen und die Krallen zu zeigen.

39. Simba verbietet Mietzi zu klettern und auf die Jagd zu gehen.
40. Simba verbietet, Mietzi zu locken und auf die Jagd zu gehen.

41. Kevin hilft Karsten zu malern und schön sauber zu machen.
42. Kevin hilft, Karsten zu kämmen und schön sauber zu machen.

43. Karsten hilft Lena zu puzzlen und das Rätsel zu lösen.
44. Karsten hilft, Lena zu kitzeln und das Rätsel zu lösen.

45. Lena lehrt Kevin zu rechnen und die Zahlen zu schreiben.
46. Lena lehrt, Kevin zu fragen und die Zahlen zu schreiben.

47. Mietzi lehrt Simba zu schnurren und die Ohren zu pflegen.
48. Mietzi lehrt, Simba zu kraulen und die Ohren zu pflegen.

49. Kevin verspricht Lena zu spucken und ganz schnell abzuhauen.
50. Kevin verspricht, Lena zu kratzen und ganz schnell abzuhauen.

51. Lena verspricht Kevin zu heulen und danach frech zu grinsen.
52. Lena verspricht, Kevin zu zwicken und danach frech zu grinsen.

53. Peter bittet Anna zu hopsen und ein Liedchen zu pfeifen.
54. Peter bittet, Anna zu loben und ein Liedchen zu pfeifen.

55. Anna bittet Peter zu klatschen und sehr laut rumzubrüllen.
56. Anna bittet, Peter zu schimpfen und sehr laut rumzubrüllen.

57. Anna erlaubt Lena zu quatschen und ein Liedchen zu trällern.
58. Anna erlaubt, Lena zu schminken und ein Liedchen zu trällern.

59. Stella erlaubt Bello zu kämpfen und dann lange zu schmusen.
60. Stella erlaubt, Bello zu kraulen und dann lange zu schmusen.

61. Stella verbietet Bello zu knurren und das Kind umzuwerfen.
62. Stella verbietet, Bello zu quälen und das Kind umzuwerfen.

63. Bello verbietet Stella zu jaulen und ganz laut rumzuwinseln.
64. Bello verbietet, Stella zu schubsen und ganz laut rumzuwinseln.

65. Lena hilft Kevin zu quengeln und ganz laut rumzuheulen.
66. Lena hilft, Kevin zu tadeln und ganz laut rumzuheulen.

67. Kevin hilft Peter zu toben und danach rumzukichern.
68. Kevin hilft, Peter zu knebeln und danach rumzukichern.

69. Peter lehrt Anna zu lügen und die Taschen zu stehlen.
70. Peter lehrt, Anna zu stossen und die Taschen zu stehlen.

71. Bello lehrt Stella zu schnüffeln und die Gefahr zu meiden.
72. Bello lehrt, Stella zu warnen und die Gefahr zu meiden.

73. Kevin verspricht Tina zu beten und ganz leise zu spielen.
74. Kevin verspricht, Tina zu mögen und ganz leise zu spielen.

75. Lena verspricht Maxe zu bleiben und die Erbsen zu essen.
76. Lena verspricht, Maxe zu knuddeln und die Erbsen zu essen.

77. Thomas bittet Kevin zu rülpsen und die andern zu stören.
78. Thomas bittet, Kevin zu ziepen und die andern zu stören.

79. Maxe bittet Tina zu lächeln und das Lied mitzusingen.
80. Maxe bittet, Tina zu grüssen und das Lied mitzusingen.

81. Gudrun erlaubt Thomas zu rutschen und sehr lange zu schaukeln.
82. Gudrun erlaubt, Thomas zu wiegen und sehr lange zu schaukeln.

83. Tina erlaubt Kevin zu schmatzen und danach loszulachen.
84. Tina erlaubt, Kevin zu pieksen und danach loszulachen.

85. Tina verbietet Lena zu jammern und den Hund anzuschreien.
86. Tina verbietet, Lena zu kneifen und den Hund anzuschreien.

87. Lena verbietet Thomas zu motzen und später wegzulaufen.
88. Lena verbietet, Thomas zu reizen und später wegzulaufen.

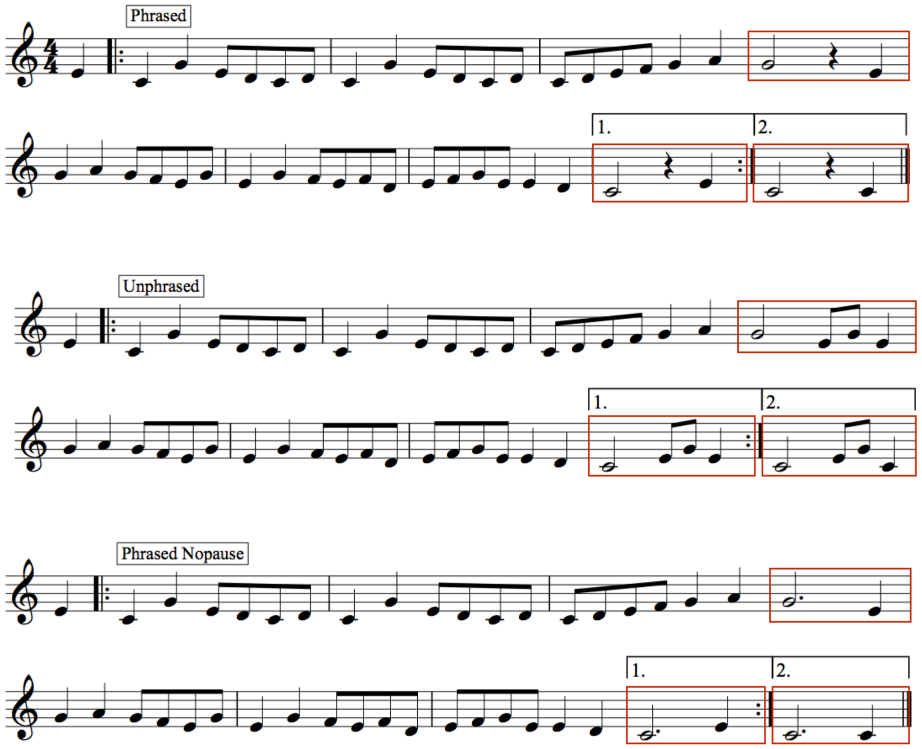
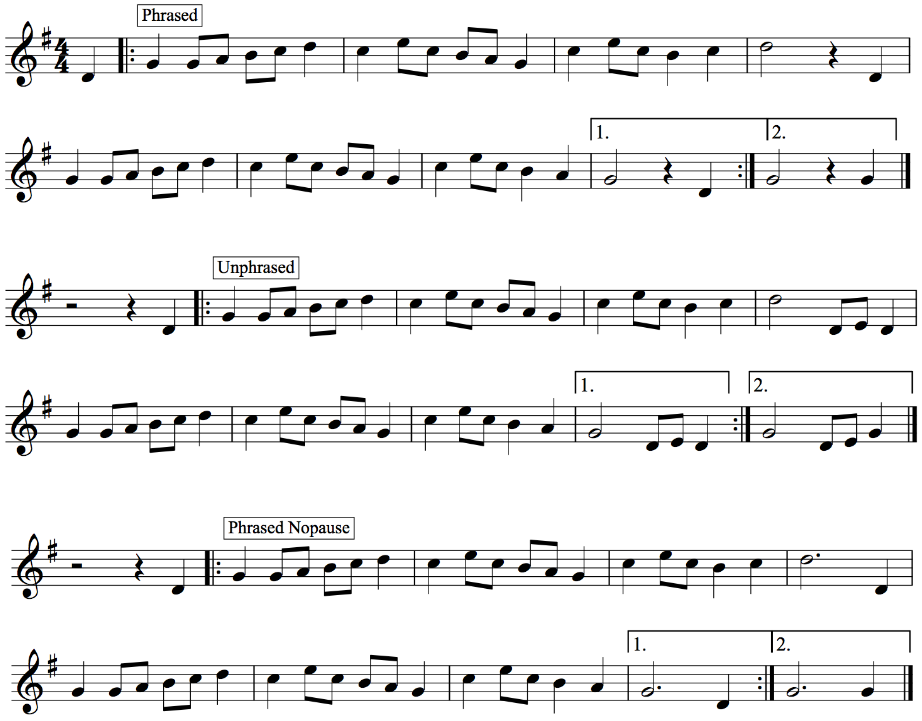
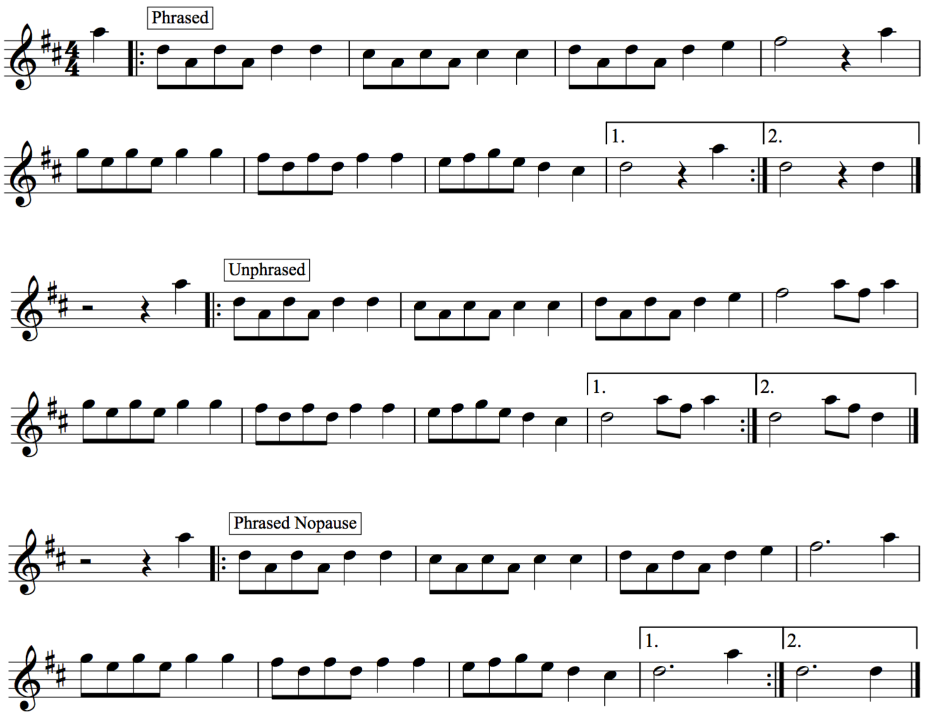
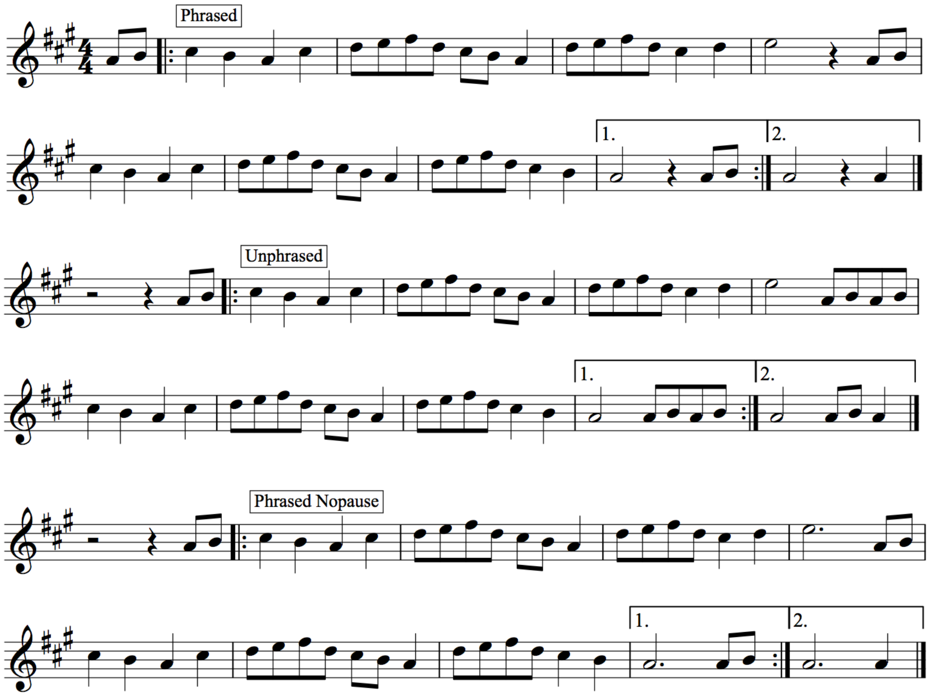
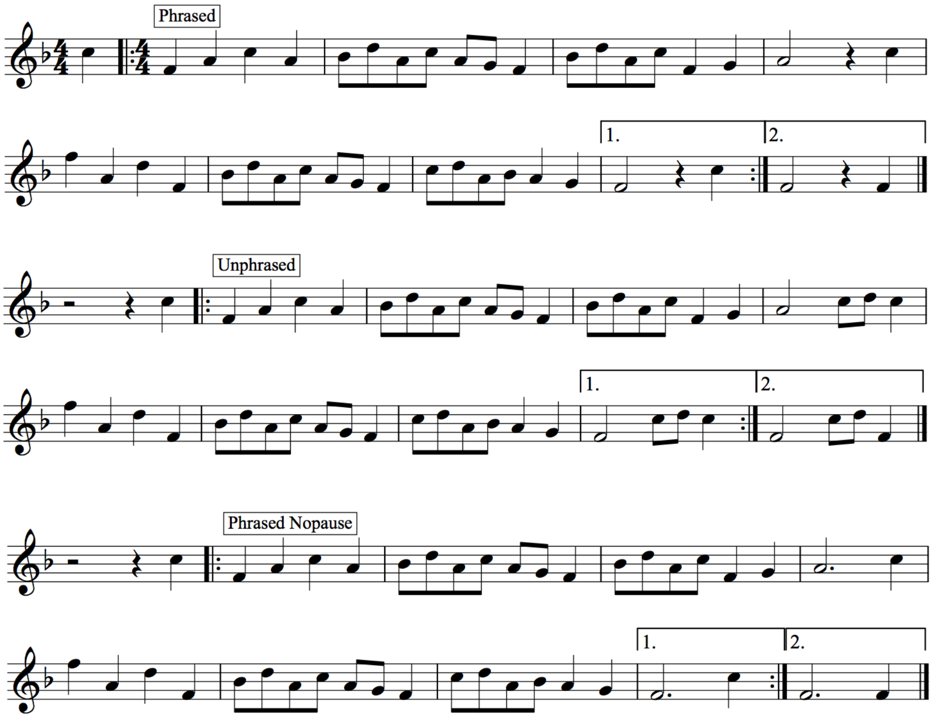
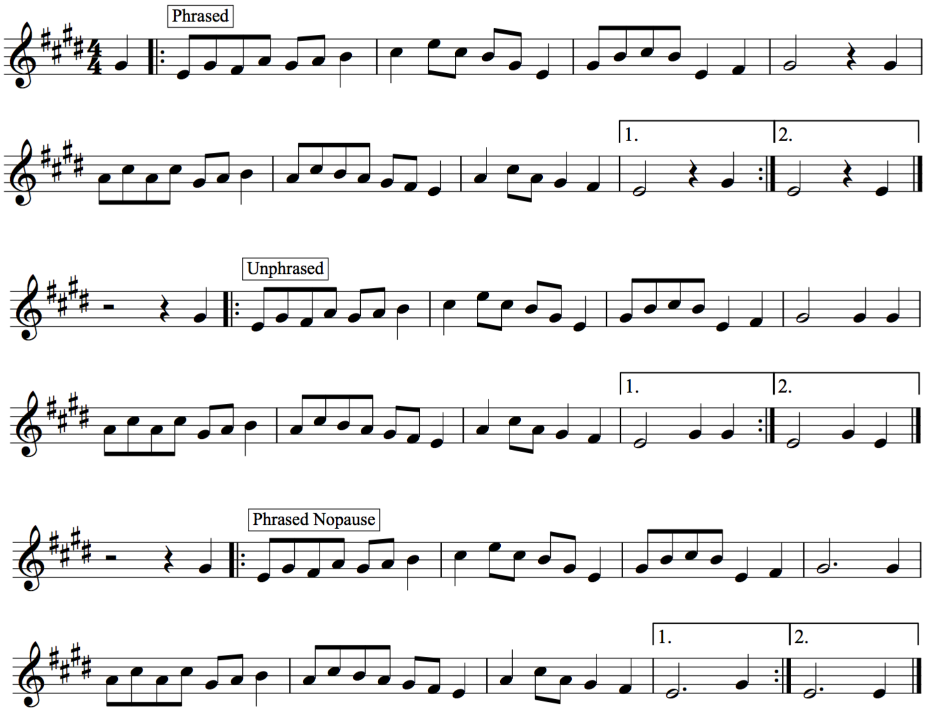
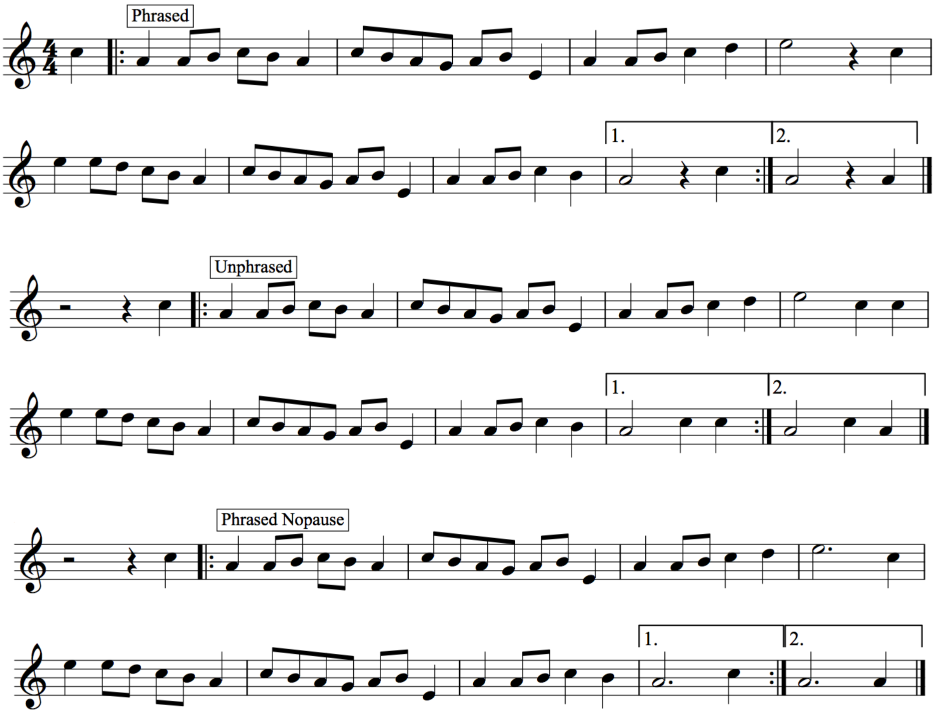
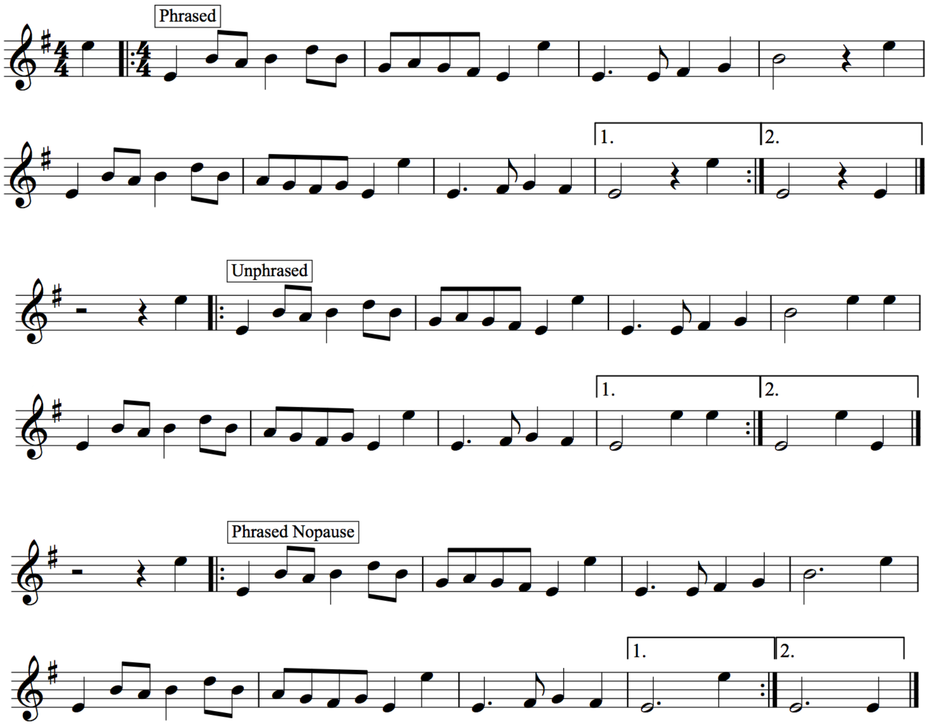
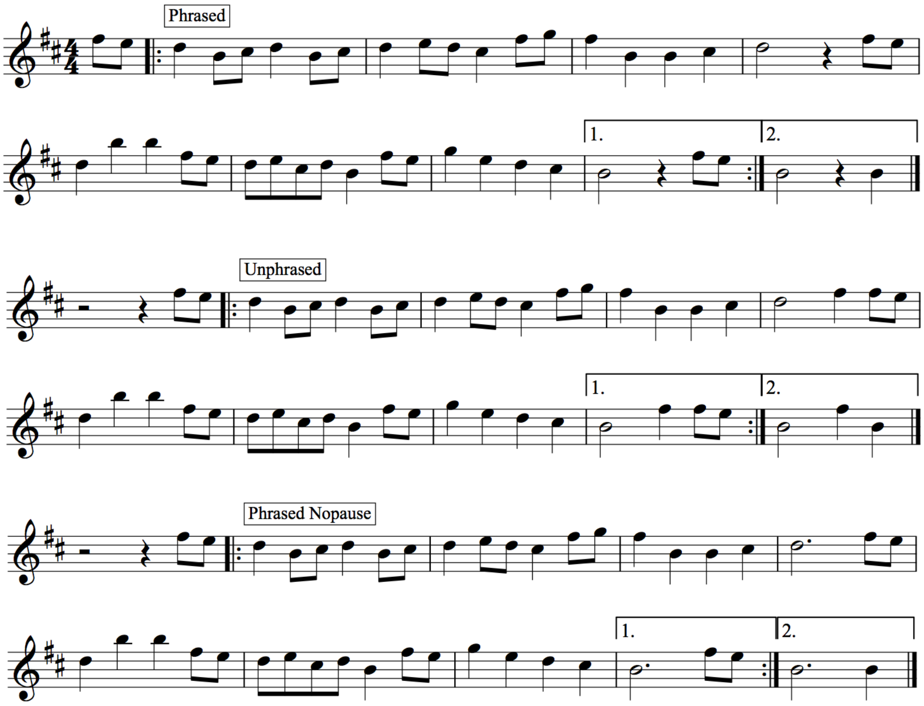
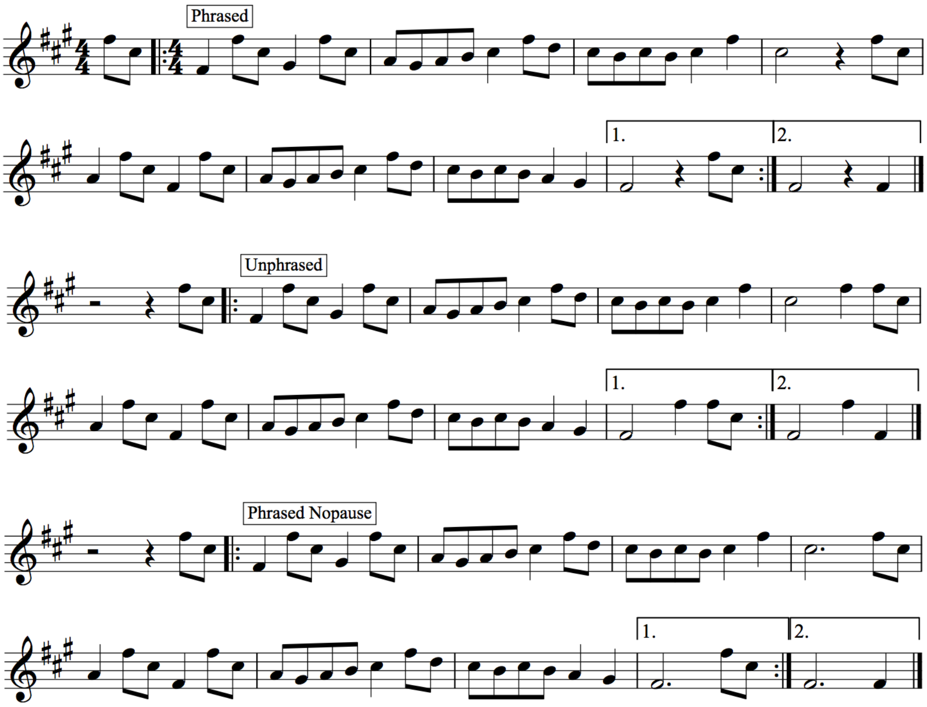
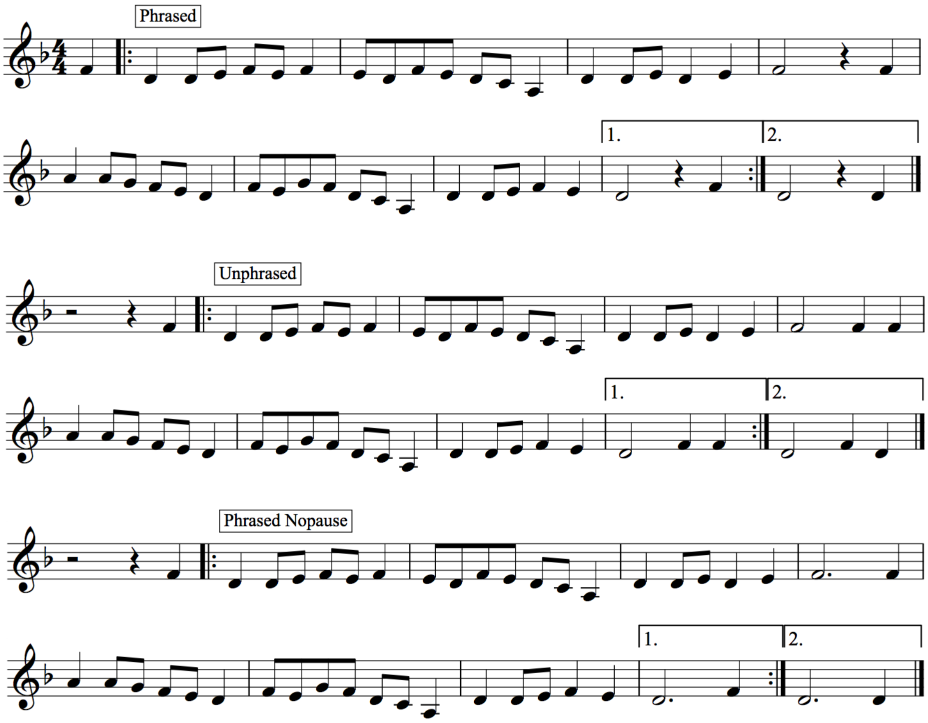
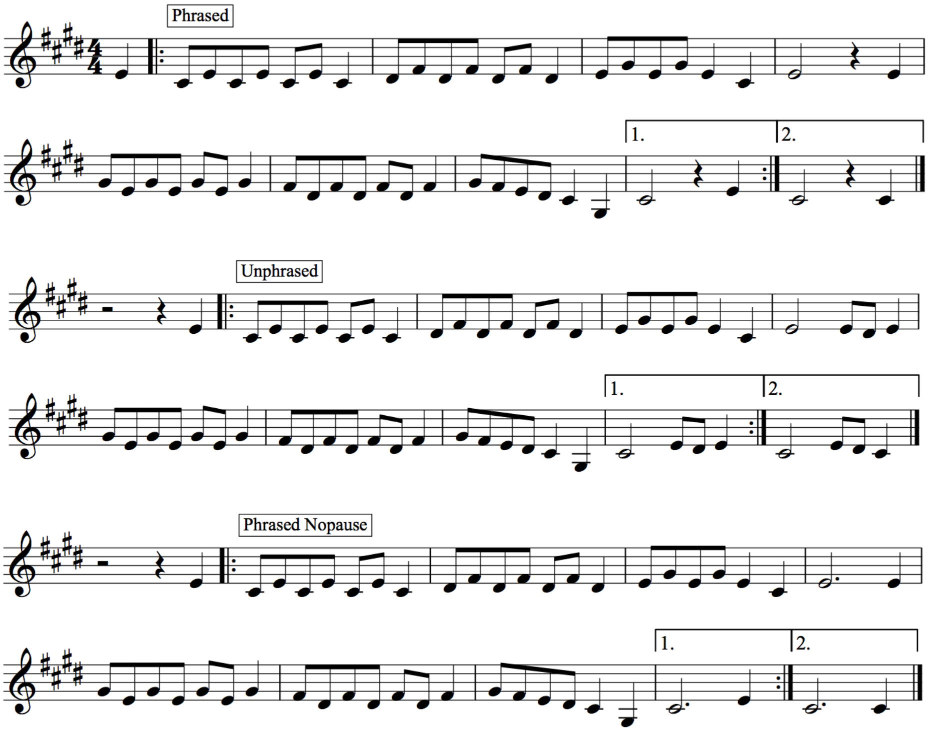
89. Maxe hilft Gudrun zu nageln und das Bild aufzuhängen.
90. Maxe hilft, Gudrun zu malen und das Bild aufzuhängen.

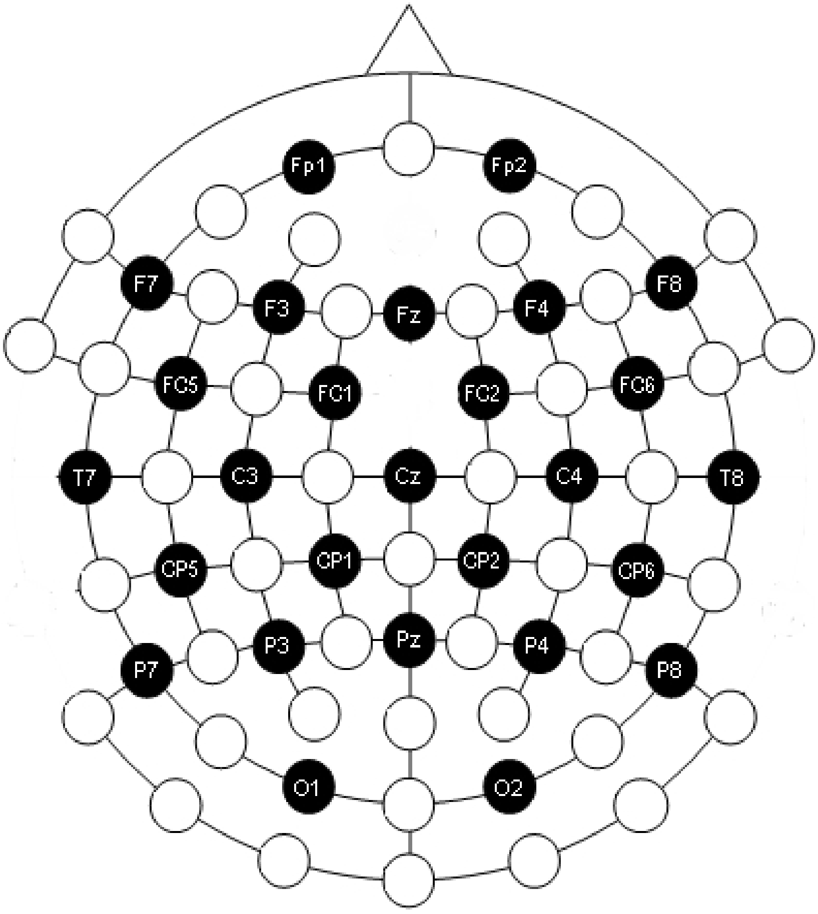
91. Kevin hilft Gudrun zu schummeln und alle auszulachen.
92. Kevin hilft, Gudrun zu necken und alle auszulachen.

93. Thomas lehrt Lena zu hocken und ganz schnell wegzuhüpfen.
94. Thomas lehrt, Lena zu retten und ganz schnell wegzuhüpfen.

95. Gudrun lehrt Maxe zu poltern und das Bad naSSzuspritzen.
96. Gudrun lehrt, Maxe zu duschen und das Bad naSSzuspritzen.

7.2 Appendix II: Music Stimuli

Stimulus 1

Stimulus 2

Stimulus 3

Stimulus 4

Stimulus 5

Stimulus 6

Stimulus 7

Stimulus 8

Stimulus 9

Stimulus 10

Stimulus 11

Stimulus 12

7.3 Appendix III: Electrode Configuration



Biography

John DePriest grew up in Columbus, Mississippi, and Alexandria, Virginia, and developed an interest in language and music from an early age. He began learning spanish while learning to speak English, and began playing cello at the age of 4, and piano at 8. In high school, John took every class offered in Spanish and German, so that the German teacher had to develop a new level every year. During this time he also began playing banjo and guitar, instruments that have defined his life ever since. This love of music brought him to Nashville, where he played music in numerous bands when not absorbed in his studies of German and Psychology at Belmont University. This led to the opportunity to spend a semester studying at the TU in Dresden, becoming more familiar with the German schools of psychology and neuroscience. After graduation, John pursued a career in beer brewing, which brought him to south Mississippi to work for Lazy Magnolia, while the effects of hurricane Katrina were still daily features of life across the region. Not long after, he decided to change directions and moved to New Orleans to apply to graduate school in the Linguistics Program of Tulane University. Studying in the linguistics program exposed John to endangered languages, acoustic phonetics, musical syntax, and EEG methodology. He synthesized several of these diverse interests into the study of language and musical phrase-boundary processing mechanisms, eventually moving to Berlin to design and implement a study involving neuro-typical and autism spectrum participants at the Freie Universität. After returning to New Orleans, John taught classes on Language & Memory, and Language & Music at Tulane, while serving as associate editor of the new Tulane linguistics working papers journal, *Fleur de Ling.* When not conducting or analyzing research, John plays banjo in numerous bands around New Orleans, including such styles as country, bluegrass, funk, gospel, and rock and roll. In the future, John is interested in conducting music and language neuroscience research, with particular interests in hierarchical musical syntax, harmonic closure and cadence, prosodic features of discourse, distinctions between prosody and melody, and phrase-boundary processing.