9-2016

An ERP Study of Sensory-Linguistic Processing in the Context of ASD Research

Larissa R. Miller
The Graduate Center, City University of New York

AN ERP STUDY OF SENSORY-LINGUISTIC PROCESSING
iN THE CONTEXT OF ASD RESEARCH

by

LARISSA MILLER-THESING

A master’s thesis submitted to the Graduate Faculty in Linguistics in partial fulfillment of

the requirements for the degree of Master of Arts,

The City University of New York

2016

AN ERP STUDY OF SENSORY-LINGUISTIC PROCESSING
iN THE CONTEXT OF ASD RESEARCH

by

LARISSA MILLER-THESING

This manuscript has been read and accepted for the Graduate Faculty in Linguistics in satisfaction of the thesis requirement for the degree of Master of Arts.

Date	**Valerie Shafer** **Thesis Advisor**
Date	**Gita Martohardjono** **Executive Officer**

THE CITY UNIVERSITY OF NEW YORK

ABSTRACT

AN ERP STUDY OF SENSORY-LINGUISTIC PROCESSING iN THE CONTEXT OF ASD RESEARCH

by

LARISSA R. MILLER-THESING

Advisor: Valerie Shafer

This thesis considers the questions of what is hindering the language development of children with Autism Spectrum Disorder (ASD). Tomasello's (2003) theory of language development by use necessitates adequate pattern recognition, intention reading and use for language development. Within these areas there are many variables where language acquisition could become derailed, from basic attention, perception and memory to higher cognitive functions, as well as processing speed and synchronization. Although there is a wide range of language abilities for children on the autism spectrum, the one consistently under developed area is pragmatics. One of the key aspects of pragmatic processing is making observations in the visual domain, while simultaneously processing speech. Research shows atypical and asynchronous audio visual processing for individuals with ASD, but more research is needed on the realtime multimodal processing requirements that are specific to typical language development in order to understand the ways in which atypical or asynchronous processing may affect language development for children with ASD.

The ERP study in this thesis tested the differences in brain responses to photographs paired with related sentences compared to responses to fixation symbols with sentences, as well as photographs with tones. Eight typical children between the ages of six and thirteen participated. An examination of ERPs at frontal and occipital sites to photograph onset resulted in a consistent double peaked response whether or not photographs were followed by sentences or tones. Responses to fixation symbols had a P1 but lacked a P2 response. Brain responses to audio onset had less consistent results across participants, but generally showed a late slow positive wave at occipital sites in conditions where photographs were available to view. On average larger responses were found to sentence onset than to tone onset.

These findings suggested that typical children's brain responses show an effect of voluntary attention to photographs, which is enhanced when prompted by the introduction of a related sentence. The more we understand about intermodal linguistic processing for typically developing children, the more we will be able to discover what has gone awry in language acquisition for children on the autism spectrum.

TABLE OF CONTENTS

LIST OF GRAPHS

Introduction

This thesis will first consider the question of what is hindering the language development of children with ASD. This will be followed by the background on a neurophysiological study that examined language processing in minimally-verbal children with ASD and which motivates the experiment reported in this thesis. In the second part, the method and results of study are reported followed by a discussion of these results in light of typical and atypical language development.

Theories of ASD

Autism Spectrum Disorder (ASD) is greatly heterogenous. Under the *Diagnostic and Statistical Manual of Mental Disorders* 5th Edn. criteria, the current diagnosis of ASD includes people with a wide range of linguistic abilities from non-verbal to excellent verbal ability, all with notable pragmatic and sensory differences. A single cause or comprehensive profile has yet to emerge.

Initially, ASD was misunderstood as primarily an intellectual disorder or as the result of detached parenting, the so called the Refrigerator Mother Theory (Leo Kanner; 1949). Growing understanding and the search for a cause for autism have led to various unifying theories. Some unifying theories categorized ASD as a conceptual deficit leading to a social emotional disorder. One unifying theory is Weak Central Coherence (Frith and Happe, 1994), where the core deficit is an inability to create a big picture context. Another is Theory of Mind (Baron-Cohen, Leslie and Frith, 1985) which characterizes autism as an inability to take another's perspective. These theories describe major challenges of people with ASD. The question here is whether they suffice as root causes for the linguistic difficulties of people with autism.

Since the symptoms of ASD include a wide range of linguistic abilities, considering the linguistic perspective while assessing a root cause may be elucidating. Can a core deficit account for why non-verbal children with ASD do not develop fluency in their native language despite adequate intellect and a sufficient language learning environment, while others excel verbally, and exhibit only pragmatic difficulties? Is having Theory of Mind or strong Central Coherence driving or largely facilitating a typical toddler's acquisition of language? To what degree can one develop phonological, syntactic, semantic or pragmatic abilities without it? How does this fit with what we understand about typical language development?

According to Tomasello's (2003) model of language acquisition, language is acquired through use by practice in the first few years of life. According to this theory, meaning is delineated by how language is used, and linguistic structure emerges through use. A child's ability to recognize patterns allows them to develop their grammar and pragmatics is developed through intention reading (Tomasello, 2003).

By applying Tomasello's (2003) model, core deficit theories for ASD may be evaluated for their ability to explain the variation in language acquisition within the population of ASD. For example, Weak Central Coherence would restrict language use and acquisition. The lack of an ability to put utterances into real world, realtime context would result in difficulties reading intentions and using pragmatics. Contexts are necessary for pattern recognition as well. Semantic difficulties would follow from a lack of referential and logical context. And, syntax and phonology require structurally contextualized words, and phonemes. However, high functioning children with ASD only have trouble using pragmatics. Phonological, syntactic, and logical semantic contexts are available to them. If Weak Central Coherence only applies to pragmatic contexts, then it does not, on its own, account for why some individuals with ASD are non-verbal. If it applies to all contexts then high functioning individuals should not have been able to develop the language skills that define them as high functioning.

Language use could also become restricted without Theory of Mind. As long as we assume intention reading depends on a Theory of Mind, then pragmatics and likely semantics with subjective referents would fall outside the scope of acquisition. While, phonology, syntax and semantics using objective referents could still be acquired through pattern recognition. Lacking a Theory of Mind alone is not enough to keep non-verbal children with ASD from developing basic linguistics structures and concrete references that could be acquired through pattern recognition.

Intense World Syndrome

One of the more recent theories for ASD is Intense World Syndrome (Markram, Rinaldi and Markram, 2007). It looks at ASD as primarily a processing problem. Although this theory is newer, considering perception and processing as the core issues for ASD is not new (Bogdashina, 2013). According to Intense World Theory, the brain develops over connectivity in local brain circuits, creating hyper-attention, hyper-perception, and hyper-memory. This leads to an overwhelming experience of a normally unthreatening environment causing disengagement and controlling coping strategies. In this theory, varying symptoms within the ASD population are explained by different combinations of brain areas being affected for individuals (Markram, Rinaldi and Markram, 2007).

Assessing Intense World Syndrome in terms of language acquisition according to Tomasello's (2003) framework of language use is not straightforward. Here, anxiety, disengagement and maladaptive behaviors could be considered the major obstacles to language use and acquisition. At the same time, hyper-attention, hyper-perception and hyper-memory may also affect pattern recognition and intention reading abilities.

A full assessment of how hyper-processing would affect pattern recognition requires a more detailed consideration of pattern recognition theories than is within the scope of this paper. But fundamentally, attention, perception and memory are essential to all pattern recognition theories. It is obvious that hypo-attention, hypo-perception and hypo-memory would be obstacles to pattern recognition, but hyper-processing by definition is not optimal and may also affect pattern recognition by introducing noise into the learning cycle.

Hypothetically, hyper-processing could also encourage strategies or mechanisms of pattern recognition that are not optimal specifically for language use. Theories of pattern recognition utilize different methods, such as, template matching, prototype matching, feature matching, structural analysis, statistical modeling, etc. Methods vary in the amount of attention, perception and memory employed (Asht, and Dass; 2012).

For example, template matching requires a large amount of memory to store individual pattern templates. The advantage of pattern recognition by template is that processing is fast for a small number of routinized patterns. The disadvantage is a less flexible system that does not handle variation well. For comparison, feature matching requires breaking patterns down into key features in order to recognize patterns. It takes more time, but less memory and is more flexible. It is possible hyper-memory may support pattern recognition by an over specified method, whereas a mechanism for pattern recognition that accounts for variation is more effective and sustainable over the course of language development.

More research is needed to discover if people with ASD have appropriately gauged attention, perception and memory to break down information on every linguistic level. Echolalia, or reciting entire scripts by memory, suggests that autistic children who exhibit this symptom are not processing linguistic information in a way that facilitates breaking down the speech stream into usable building blocks. In general, more information is needed about how hyper-processing versus hypo-processing affects pattern recognition for language acquisition.

In terms of intention reading, not enough is known about how hyper-attention, hyper-perception and hyper-memory would affect intention reading to assess Intense World Syndrome for language acquisition on the pragmatic level. But, a listener's interpretation of what is communicated in a speech event is more likely to match the speakers intent, if they share attention, perceptions and memories. More research is needed to discover if atypical processing undermines some of the shared assumptions that aid intention reading. More information is also needed about how hyper-attention, hyper-perception and hyper-memory affects the ability to hone in on important pragmatic observations during speech events.

Pragmatics and Intention Reading for Children with ASD

With respect to language acquisition, the one commonality for people with ASD as currently defined by the *Diagnostic and Statistical Manual of Mental Disorders, 5th Edition* (DSM V; 2013) is that they do not fully acquire all aspects of pragmatic language no matter their verbal ability. People with ASD do not adhere to most of Grice's (1975) Cooperative Principles. The one Maxim they do follow is Grice's Maxim of Quality, in that they generally tell the truth, possibly at a cost to the other Maxims. Frequently, they do not adhere to the Maxim of Quantity, speaking either too much or not enough. They often do not adhere to the Maxim of Relevance, not always responding appropriately to questions or staying on topic. And, they do not adhere to the Maxim of Manner with respect to being succinct and orderly, taking turns, not talking over people, responding within an expected timeframe, etc. Flouting Maxim's is not a willful act of non-cooperation. It is a symptom.

In order to understand the lack of pragmatic language acquisition for children with ASD, more needs to be discovered about how intentions are read, since there is no consensus about exactly how typical individuals read intentions. What is certain is that intention reading cannot require mind reading or typical individuals would also fail at the task.

Although the case of sarcasm, jokes and hyperbole are not the average case of intention reading, it is an area that has received attention. According to models explaining sarcasm, intention reading includes reasoning in which a bridge is made from an observable speech event to an unobservable intention. Whether intention reading is positioned as induction or abduction, the process of learning to read intention relies on the observation of pragmatic cues and shared assumptions about real experience.

According to Grice (1975) sarcasm, jokes and hyperbole require the minimal recognition that these types of utterances are marked as other than literal or true statements, which is the assumed default intention. Speakers use pragmatic cues, such as prosody, facial expressions, gestures, body language etc. to indicate that an utterance has diverged from the default statement pattern. Without noticeable pragmatic cues a listener may interpret that the speaker intended the utterance to be taken as a true statement (Grice, 1975). In this case intention reading is accomplished through complex, realtime, multisensory pattern recognition.

Other theories have countered or modified Grice's model, such as Mention Theory (Jorgensen, Miller and Sperber,1984) and Pretense Theory (Clark and Gerrig, 1984). In Pretense Theory for example, there is no default interpretation. The speaker expresses their attitude toward a situation through pretending to be ignorant of the situation. The pretense is perpetrated with pragmatic cues, which are often exaggerated. The real situation and the speaker's intention is considered to be obvious for the listener. In this case, intention reading is accomplished through the listener's best guess.

Here is an example that Clark and Gerrig (1984) take from Jorgensen et al., (1984) and reinterpret using Pretense Theory: *A woman says: Trust the Weather Bureau! See what lovely weather it is: rain, rain, rain.* (p. 122). In this example it is raining and the the speaker pretends not see the rain. This type of pretense is meant to be discovered and is accompanied by overacting. A listener who sees that it is obviously raining and notices the exaggerated tone, facial expressions, gestures etc. will understand that the speaker means to make fun of the Weather Bureau's mistake. A listener who misses the cues and does not see that it is raining will miss the intended meaning.

In both Mention Theory and Pretense Theory, an important foundation for learning how to read intention is observation. Pragmatic cues are communicated in the moment by the speaker, with the assumption of shared processing abilities. Speakers assume listeners will notice both visual and auditory cues, and take into account all the pertinent information, as well as process what is said, and they assume discourse will move along at a typical pace. If learning to read intention depends upon realtime, simultaneous, multisensory observation then processing that is atypical in range, or asynchronous to the norm are plausible obstacles to pragmatic use and acquisition.

Language Acquisition and Entrainment Rhythms

Timely processing is needed for intention reading and pragmatic use and it is also an essential aspect of real-time pattern recognition. According to Mari Jones (1976) sense of time and rhythm are fundamental to attention, perception and memory. In her article on dynamic attending, Jones points out that in the case of real-time pattern recognition, it is not only important to know what to pay attention to, but when to pay attention. Missing required attentional moments will affect what is perceived and remembered and therefore will affect pattern recognition overall.

According to Barnes and Jones (2000) knowing when to pay attention is aided by expectancy. In experiments they showed that internal rhythms naturally synch up to external timing. This entrainment causes expectancy, which guides when to pay attention.

Since the optimal state of attention is to be in synch, premature attention is no more advantageous than delayed attention for real time pattern recognition.

Entrainment is evidenced in oscillations. Studies of resting state oscillations for participants with ASD, show atypical internal rhythms without the presentation of stimuli. They have found reduced delta, theta, beta and gamma oscillations and enhanced alpha oscillations, as well as enhanced power in the left hemisphere compared to the right (Wang et al., 2013). Atypical oscillations have also been found for populations with ASD when a variety of stimuli are presented. A combined Functional Magnetic Resonance/ Electroencephalography (fMRI/EEG) study of thirteen participants with ASD and a control group was done by Jochaut et al.,(2015). This study found that “gamma and theta cortical activity do not engage synergistically in response to speech.” (Jochaut et al., 2015). Atypical gamma waves have also been found in adults with ASD when presented with faces (Sun, et al.,, 2012).

The natural process of entrainment may play a foundational role in language acquisition. Language processing requires auditory processing on different interval scales. As people age, they entrain to progressively longer time scales, although not steadily (McAuley et al., 2006). For example, there is a jump at age eight which corresponds to extended duration of attentional abilities in general and changes in motor control (McAuley et al., 2006). McAuley et. al. suggest that the course of attentional entrainment is geared toward facilitating language acquisition, capturing auditory patterns of the shortest duration first and moving to longer patterns with increasing age.

The shortest scale linguistic information is the phoneme. Auditory studies have shown that well before gathering the phonemic attributes of their language, newborns are already attuned to the rhythms of their native language environment (Kuhl, 2004). When infants begin babbling sometime between five and ten months, they use the prosodic rhythm of their native language (Whalen, Levitt and Wang, 1991) even though they have not narrowed down to the phonemic categories of their own language until around age one. This suggests that the timing of when to find salient features for pattern recognition is acquired before phonemic category itself in the course of language acquisition.

As with auditory language, visual expressions occur at different time scales. Normal facial expressions last anywhere from 500 milliseconds to 4 seconds in length and are often reiterating following the rhythm of speech. Physical expressions carry important information pertinent to speech events.

In terms of pattern recognition, an analogy may be made to phonology in that physical expressions are comprised of various combinations of features, such as eyebrow position, eye shape, mouth position, head position, hand position and body position. Plus, the size and duration of a physical expression will affect salience. For example, in addition to normal facial expressions, there are micro expressions which are subtle and last around 50 ms. As opposed to normal expressions, micro expressions are considered to be below the threshold of the average listener's conscious perception and carry information that was intended to be concealed by the speaker (Haggard and Isaacs, 1966).

Visual entrainment can be seen in the form of a listener's eye blink synchronization to speaker eye blinks. Eye blink synchronization only occurs in adults when they are able to view the speakers whole face. Entrainment is lost in adults when listeners are restricted to viewing either the mouth or the eyes alone. Adult listeners with ASD do not synchronize eye blinks with speakers under any conditions even when the whole face is in view (Nakano, Kato, and Kitazawa, 2011).

In adults, visual entrainment synchronizes to prosody with eye blinks occurring most frequently during pauses in the speech stream (Nakano and Kitazawa, 2010). Visual expectation has been seen in typical infants as young as 3.5 months old when they were shown images at regular intervals as opposed to irregular intervals (Haith, Hazan and Goodman 1988). While looking at speaking faces, infants fixate more on eyes before 4 months of age, then switch to look more at mouths from 4-8 months and then switch back to eyes at around 12 months old (Lewkowicz and Hansen-Tift 2012). It is possible that this pattern shows scales of entrainment as well as audio visual binding, beginning with eye blinks and prosodic rhythms pre-verbally, then switching to mouth position for phoneme acquisition, returning more to the eyes to take in facial expressions corresponding to words and longer utterances, and ultimately progressing toward perceptually cohesive language fluency on all linguistic levels.

At the start of language acquisition, before a baby knows what to expect, knowing when to expect the next bundle of pertinent information is basic to pattern recognition. The natural process of audio and visual entrainment as a mechanism for the earliest foundation of language acquisition may explain why infants are capable of attending to speech events without fully developed executive functioning. And, the developmental sequence of entraining to progressively longer patterns in the communicative environment may allow for tuning out information from subsequent levels before foundational shorter scale patterns have been acquired.

Language acquisition has long been understood as an inherently multisensory, temporally located endeavor. In addition to the audio visual research noted here, there is a body of research regarding the connection of motor control and the development of speech production and perception. Much less studied is the role of proprioception to pragmatic acquisition.

Hypothetically the inability to naturally entrain to the typical rhythms of multisensory speech events, not just auditory speech, could disrupt language acquisition to varying degrees. 1) By the degree to which internal rhythms are off from external speech event rhythms. 2) If entrainment is possible for some time scales but not others, or in a developmentally inappropriate order. 3) If atypical or asynchronous sensory binding interferes with coordinated entrainment, or allows entrainment to some modes and not others.

So far, the audio visual research points toward a course of language development for ASD that is altered from the norm minimally by skipped entrainment to short scale eye blinks and prosody. And, where there is verbal ability, it is acquired without well coordinated visual speech and facial expressions, (as will be further reviewed in the following section.) Potentially, non-verbal children may be unable to entrain to longer time scales in addition to shorter ones, making it difficult for them to recognize linguistic patterns that fall into these time scales as well.

Atypical Timing, and Binding for Audio Visual Processing in ASD

It is known that people with ASD are out of synch with their environments sensorily as well as socially and linguistically. They may exhibit a variety of behaviors that show their difficulty with sensory integration and regulation (Marco et al.; 2011), such as rocking, flapping, picky eating, trouble sleeping, trouble potty training, high pain tolerance, tactile defensiveness, covering ears, visual fixation etc. The list is long and varied. It can include both sensory seeking and sensory averse behaviors and is by no means consistent across individuals or even for a single individual who may have seemingly contradictory symptoms at different times. Sensory integration difficulties can present in a single sense or in any combination of seven senses, visual, auditory, olfactory, gustatory, somatosensory, vestibular and proprioceptive (Ayres; 1972).

Neurophysiological studies of children with ASD have found differences in processing for various types of sensory stimuli. Most studies have focused on tactile, audio or visual responses. Audio and visual studies have been done for simple stimuli as well as complex, social emotional and linguistic stimuli. A profile of atypical timing for audio and visual processing has emerged for ASD from these studies.

Atypical processing times have been found for children with ASD. Auditory Event Related Potential studies of basic sounds have found that children with ASD have shorter latencies for early components, showing different involuntary attention to simple auditory stimuli, such as those done by Novick (1980), Martineau et al., (1984) Oades et al., (1988), Ferri et al., (2003). In neural studies using the mismatch negativity measure (MMN) to test pitch discrimination without attention, Gomot et al., (2002, 2006) found shorter latencies and Ferri et al., (2003) found larger amplitudes for low functioning children with ASD. In an MMN study examining changes in pitch and phonemic category, Lepisto et. al. (2008) found that children with ASD had larger and faster responses than the control group to variations in both pitch and phonemic category alone, but their responses were diminished and slower for changes in phonemic category when variations in pitch were included in the stimuli, suggesting that a noisy context slows speech pattern recognition for children with ASD and that they have trouble identifying or inhibiting extraneous sensory information.

Studies have also shown slower processing of visual stimuli for individuals on the autism spectrum. For example, Townsend et al., (2001) found that children with ASD were less accurate and had slower response times for a visual spacial attention task than a control group and that response times for peripheral vision was slower than for central locations. This study also showed altered P3b responses, which are associated with orienting visual attention and event processing. They concluded that children with autism were delayed in updating visual contexts in working memory and had difficulty orienting their visual attention, as well as in processing latent information (Townsend et al., 2001).

One audio visual ERP study included an emotion recognition task. Lerner et al., (2013) looked at visual processing of faces and audio processing of voices. They found that information processing speed correlated with emotion recognition ability and there was a correlation between early component latencies and later component amplitude across modalities.

Not many neurophysiological linguistic studies have been done with combined audio visual stimuli for participants who have ASD. A rare example is a study by Cantiani et al., (2016). They found atypical responses for minimally verbal children with ASD compared to neurotypical children using a picture-word matching paradigm. Responses were slower for both spoken words and for pictures and significant differences were found for lexical-semantic processing at the higher order level for children with ASD (Cantiani, et al., 2016).

Behavioral studies of visual-speech integration have found that children on the autism spectrum have atypical responses to asynchronous audio visual events. Bebko et al., (2006) showed that children with ASD had no preference for synchronized voice and lip movements, as opposed to out of sync voice and lip movements. In addition, the McGurk effect was less prevalent for children with ASD according to De Gelder et al., (1991) and Williams et al., (2004). Iarocci et al., (2010) found that children with ASD reported audio stimuli more than visual stimuli. And, Smith and Bennetto (2007) found that in an environment with background noise, lip reading did help children with autism hear speech better than without, but not to the same degree as it did typical children.

Another behavioral audio visual integration study by Donahue et al., (2012) used a simultaneity judgement task. The stimuli presented were a tone and a black and white checkerboard pattern. They presented audio and visual stimuli with varying stimulus onset asynchronies (in ms: -300, -250, -200, -150, -100, -50, 0, 50, 100, 150, 200, 250, 300). Negative onsets were audio first, positive were visual first, 0 is perfectly synchronized onset of both. Participants who averaged more negative were considered audio biased and those who were more positive, visually biased.

Donahue et al., (2012) explain that, “The typical window of integration [which allows for binding different stimuli into a cohesive event] ranges from -150 to 150 ms, so subjectively, approximately half of the SOA’s [Stimulus Onset Asynchronies] should be perceived as simultaneous presentations even though only one is objectively simultaneous." (Donahue et al., 2012). They note people with ASD have longer binding windows as per Foss-Feig et al., (2011).

In the results of Donahue et al., (2012) more people considered visual first presentation as simultaneous. And, there was a correlation between those who were biased towards audio first to having higher scores on their Autism Spectrum Quotient questionnaire (higher score = more autistic symptoms). The more self reported autistic symptoms the larger the difference in multisensory integration.

In bimodal conditions studies have shown that audio and visual information do not have an equal contribution to integrated perceptions for typical populations (Hecht and Reiner 2009). In conditions where simultaneous audio and visual information conflict, visual perception takes precedence over auditory perception (Hecht and Reiner 2009). In contrast people with ASD show auditory dominance in their McGurk responses, as well as faster auditory processing speeds compared to their visual processing speeds. Auditory dominance may contribute to visual expressions being missed by people with ASD.

The Donahue et al. (2012) research suggests that visual dominance is related to appropriate binding windows and timely overall audio visual processing. Whereas, auditory dominance, long binding windows and slower overall audio visual processing are hallmarks of ASD. Whether or not the level of language acquisition in ASD correlates to the degree of asynchrony between audio visual processing for an individual, or for the overall speed of combined audiovisual processing compared to the norm remains to be determined.

Using ERPs to Study Language Development in ASD

Interacting with both sensory and linguistic environments in atypical ways are hallmarks of ASD. Online processing of sensory and linguistic stimuli can be studied with ERP's. ERP's reflect time locked brain responses to stimuli in a controlled environment. Aspects of real time language processing, such as, attention, perception and memory can be studied, as well as speed and timing. Entrainment rhythms may be investigated with oscillations. In general ERP's may be used to shed light on multimodal interactions and neural coordination for a better understanding of how processing speed and sensory integration relate to pattern recognition and intention reading abilities needed for language acquisition.

In addition, issues specific to the ASD population can be studied. The question of a core deficit may be addressed by studying individuals with a range of verbal abilities. A passive condition, requiring no behavioral response can be used for members of the population that may have severe task performance anxiety. And, individuals who are nonverbal may be studied to see if their brain responses show some latent linguistic knowledge.

Previous ASD Syntax Study

The current ERP study follows a previous study of children diagnosed with Autism Spectrum Disorder, who had minimal language functioning of approximately five words or less. The participants included one child with good verbal ability, whose diagnosis at the time of testing was PDD-NOS. An age matched typically developing control group also participated.

Materials for the previous study included 56 pairs of simple sentences and photographs. The photographs were of animals and the sentences described actions typical to the animals represented in the photographs, for eg. “The frog Jumps.” Sentences were either grammatical or ungrammatical with respect to number agreement between noun and verb. Onset of the sentence was 300 ms after the onset of the photograph, which remained visible during audio presentation.

Responses to stimuli from the onset of photographs showed that several children in the group with ASD had brain responses that started similarly to the control group, then quickly diverged from the typical responses. The typical group showed a sharp early positive peak over occipital sites, followed by a large slow positivity. The children with ASD showed the early positive peak, but had either absent or greatly attenuated responses for the second occipital positivity.

This observation led to the following questions; How does the condition of visual images and overlapping sentences affect participant brain responses? Are the children with ASD showing differences in individual modality processing or is sensory integration out of synch? Is their processing too slow to keep up? Are they not showing attentive expectancy, or are they anticipating a verbal sentence and getting overwhelmed and disengaging?

Current Experiment Methods

This follow-up study was designed to better understand the effect of bimodal linguistic processing on the brain responses of children. The current study focuses on typical children. This study uses the materials (sentences and photograph pairs) from the previous ASD syntax study.

Design

This experiment isolates sentences from photographs in two ways in order to clarify the unimodal contribution or bimodal interaction within brain responses. First, by distancing visual and audio responses from one another by increasing the onset between photo and sentence from 300 ms to 1000 ms in a photo plus sentence condition. Second, by creating two additional conditions. 1) photograph plus tone and 2) fixation plus sentence. In all three conditions, the longer onset timing was used and visual stimuli remained on during the presentation of audio stimuli.

In the condition where the photographs were replaced by a fixation symbol, a cross was presented 80% of the time and the “x” was presented 20% of the time. In the condition where sentences were replaced by tones, a 300 Hz tone was presented 80% of the time while a 400 Hz tone was presented 20% of the time. These two different frequencies were presented to allow for evaluation of attention to the target in a future study, but are not examined in the current study. The sentences were recorded natural speech, ranging in length from 1229 ms to 1909 ms. The length of sentences were analyzed as falling into natural clusters. A representative tone length was chosen for each cluster with the number of representative tones varying to match the distribution of the number of clusters.

<u>Participants</u>

Participants were four typical adults and eight typical children age six to thirteen years of age. One child was left handed. Seven children participated in the Photograph /Sentence condition. Four children participated in the Fixation /Sentence condition, and the remaining four participated in the Photograph /Tone condition.

<u>Procedure</u>

Participants sat in front of a computer screen while photographs or fixation symbols were presented and sentences or tones were heard simultaneously through two speakers located to the left and right of the computer monitor. The EEG/ERP data were recorded on an Electrical Geodesics Inc. system at 250Hz with a 65-electrode sensor and were referenced to Cz. Online filter was set at 0.1-100Hz.

<u>Analysis</u>

Offline processing included filtering: bandpass 0.3 to 30Hz and artifact rejection +/- 140 μV. Data were epoched from 200 ms before the stimulus onset (photo or sentence onsets) and baseline corrected on the 200 ms prestimulus interval. Then, epochs were checked for artifact using a criterion of +/- 140 μV. Epochs were marked as bad if more than 10% of the channels were bad. Channels were marked as bad if they were bad on more than 40% of the epochs. Bad channels were replaced by spline interpolation. After artifact reject, epochs were averaged for each participant and re-referenced to an average reference.

Comparisons of conditions from both visual and audio onset are presented. One participant in the group who participated in the Photograph /Tone condition was not able to complete the Photograph /Sentence condition. This participant's data could not be used in comparisons of conditions that included the Photograph /Tone condition.

Hypotheses

1) A difference in brain responses was expected for the different types of audio and visual stimuli used, photographs versus fixations and sentences versus tones. 2) Brain responses to the onset of the photographs were expected to show an effect from the type of auditory information to follow. A difference was anticipated for photographs when followed by a sentence compared to a tone. This effect could possibly reflect expectancy or processes such as, voluntary attention, analyzing, assessing, predicting, guessing, naming, committing to memory, etc. 3) Brain responses to the auditory information were expected to show effects from the audio visual overlap of photographs and sentences. A difference in brain responses was anticipated for the Sentence /Photograph condition compared to the Sentence /Fixation condition, due to integration of sentence and photograph or a re-assessment of photographs starting from sentence onset.

Results

<u>Photograph /Sentence Condition</u>

Figure 1 and Figure 2 show averages for seven participants in the Photograph / Sentence condition for ten sites from across the array (Fz, Cz, Oz, O1 O2, FC5, FC6, LM, LR). Figure 1 starts from photograph onset and Figure 2 starts from sentence onset, 1000ms after photo onset. Overall responses to sentences are much smaller than responses to photographs. For these ten sites, amplitudes stay between +/- 6 μV for photographs as opposed to ranging from -9 μV to 21 μV for photographs. Error bars for these graphs indicate standard error of the mean.

The pattern of brain responses to photographs shows two peaks (Figure 1). P1 (100 - 200 ms) ranges from approximately 17 μV to 21 μV for occipital sites, O1, O2, and Oz with the inversion at frontal sites, (Fz, Cz, FC5, FC6,) ranging from approximately -7 μV to -10 μV. P2 ranges from approximately 13 μV to16 μV for occipital sites with the inversion ranging from -7 μV to -10 μV for frontal sites. This second peak resolves between approximately 660 ms to 710 ms, well before sentence onset at 1000ms.

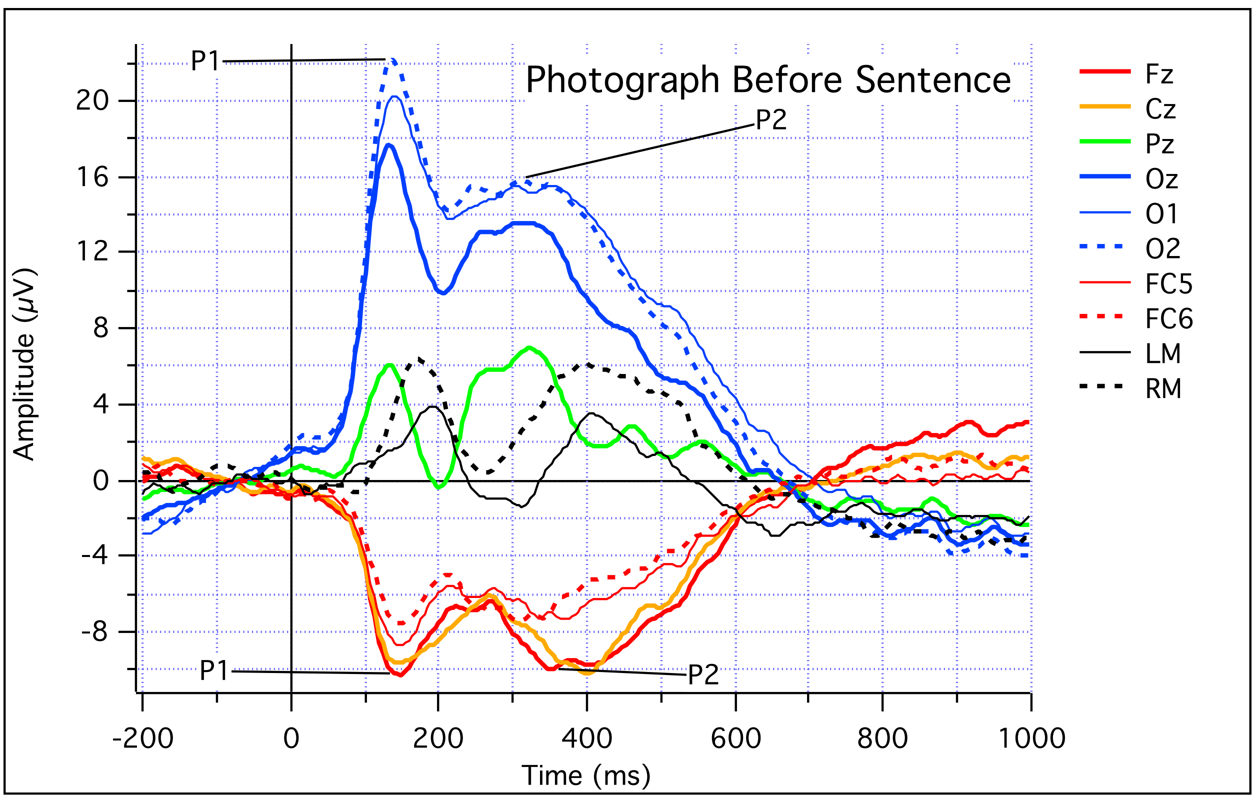

Figure 1; Averages for 7 participants in photograph/Sentence condition from photo onset; Fz(4) Cz(65) Pz (34) Oz(38) O1(37) O2(40) FC5(16) FC6(61) LM(26) RM(51).

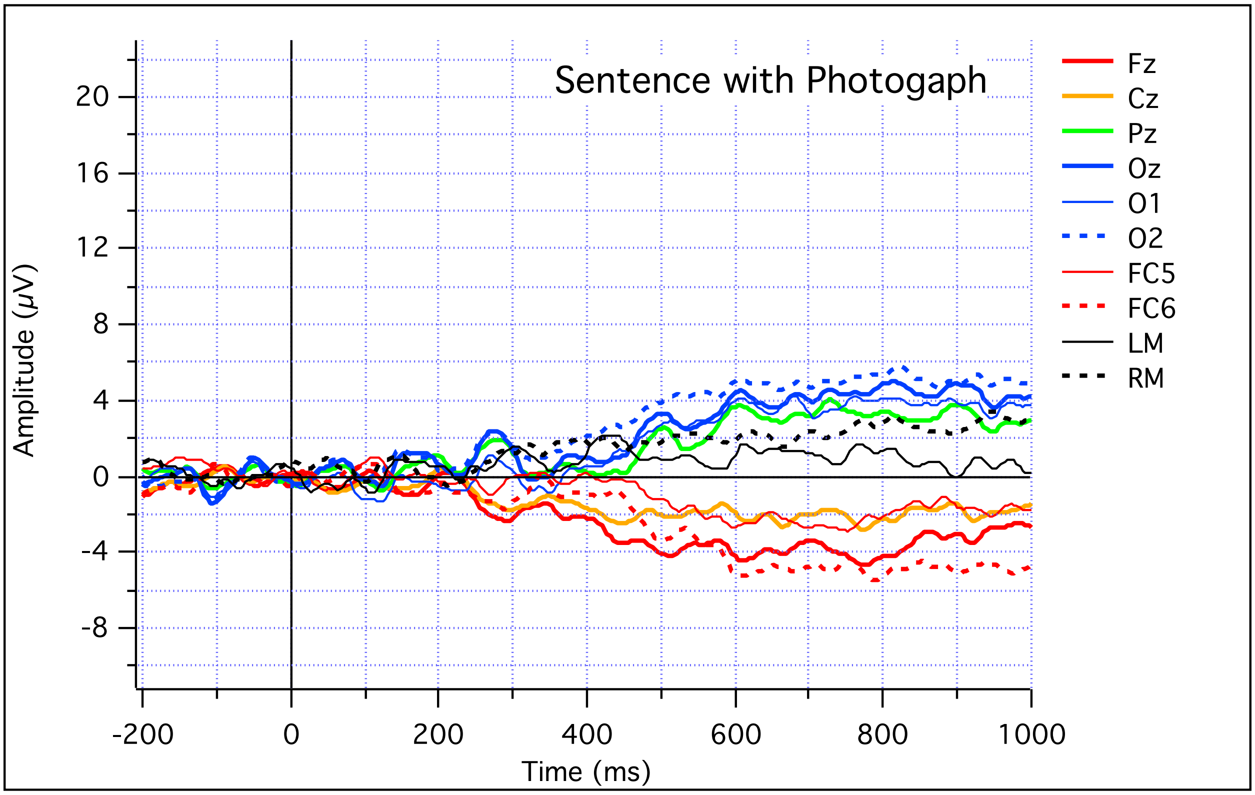

Figure 2; Averages for 7 participants in photograph/Sentence condition from sentence onset; Fz(4) Cz(65) Pz (34) Oz(38) O1(37) O2(40) FC5(16) FC6(61) LM(26) RM(51).

Condition Comparison from Visual Onset

As predicted, responses to photographs are much more robust than those to fixation symbols. Starting from visual onset, a considerable difference can be seen in comparing responses to photographs as opposed to fixation symbols (x or +) for the four participants who did both the Photograph /Sentence and the Fixation /Sentence condition (Figures 3 & 4). For occipital sites, responses to photographs for this group have amplitudes of approximately 27 μV for P1 and approximately 21 μV for P2, while responses to fixation symbols have amplitudes of approximately 10 μV for P1 and 1.5 μV for P2 (Figure 3). At frontal sites responses show the inversion pattern, here responses to photographs have approximate amplitudes of -12 μV for P1 and -12.5 μV for P2,

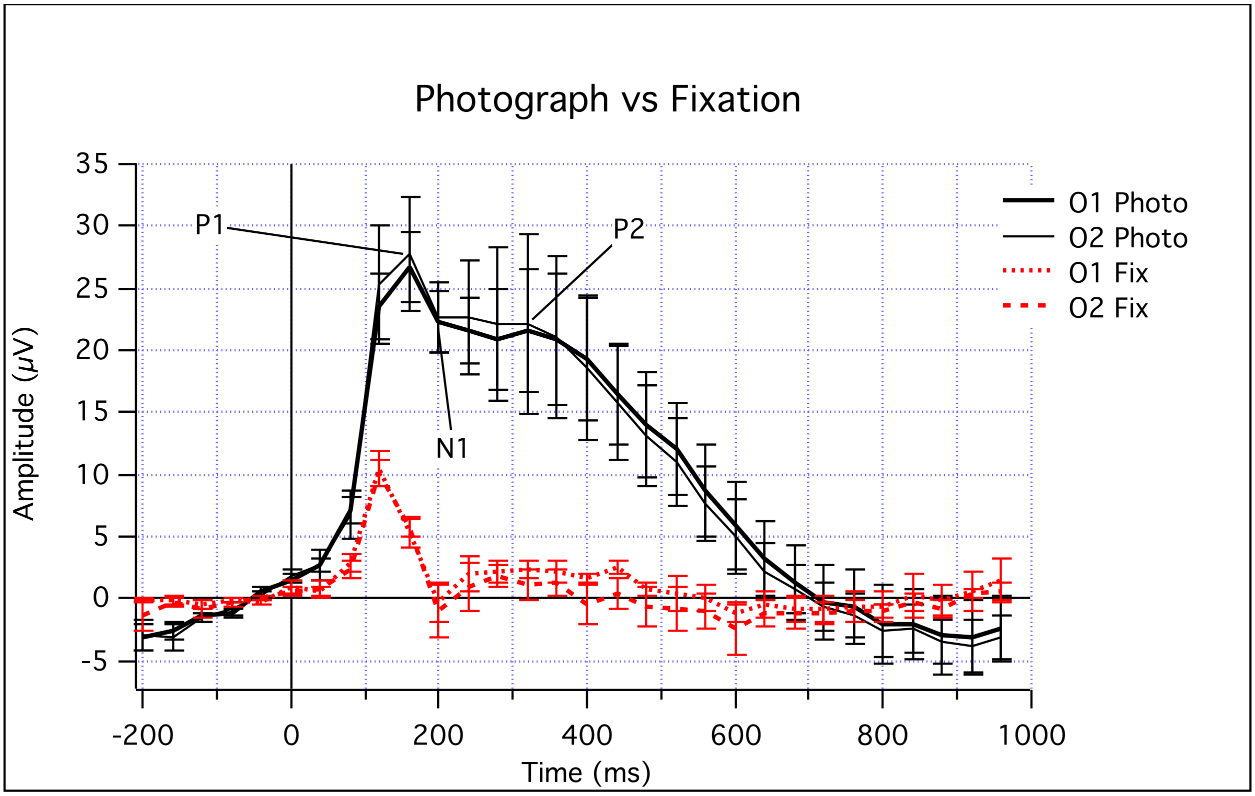

Figure 3; Averages for group of 4 participants who did Fixation/Sentence and Photograph/Sentence conditions; Comparison from visual onset; O1(37), O2(40).

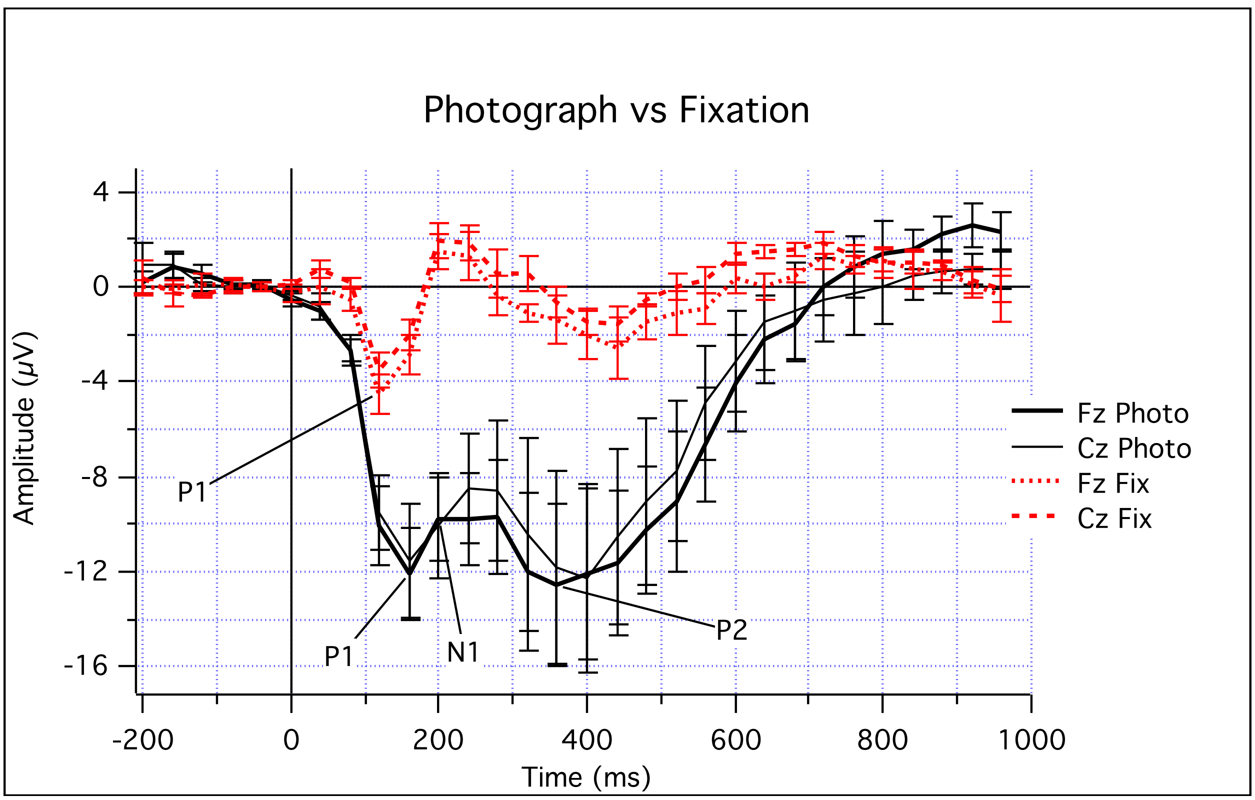

Figure 4; Averages for group of 4 participants who did Fixation/Sentence and Photograph/Sentence conditions; Comparison from visual onset; Fz(4) Cz(65).

compared to responses to fixations which have approximate amplitudes of -4.5 μV for P1 and -2.5 μV for P2 (Figure 4). Again, error bars for these graphs indicate standard error of the mean. In addition to being greater in amplitude, P2 responses for photographs are longer in duration than responses to fixation symbols, responses to photographs start at approximately 280 ms and resolve at approximately 700 ms, whereas P2 duration for fixation symbols start at approximately 200 ms and resolve at approximately 400 ms to 500 ms (Figure 3 &4).

Contrary to prediction, comparisons of responses to photographs followed by a tone, as opposed to photographs followed by a related sentence, do not show any difference of note in occipital sites or frontal sites (Figure 5 & 6). In these graphs error bars indicating the standard error of the mean overlap. These results are based on the averages of three participants, since one of the participants who did the Photograph /

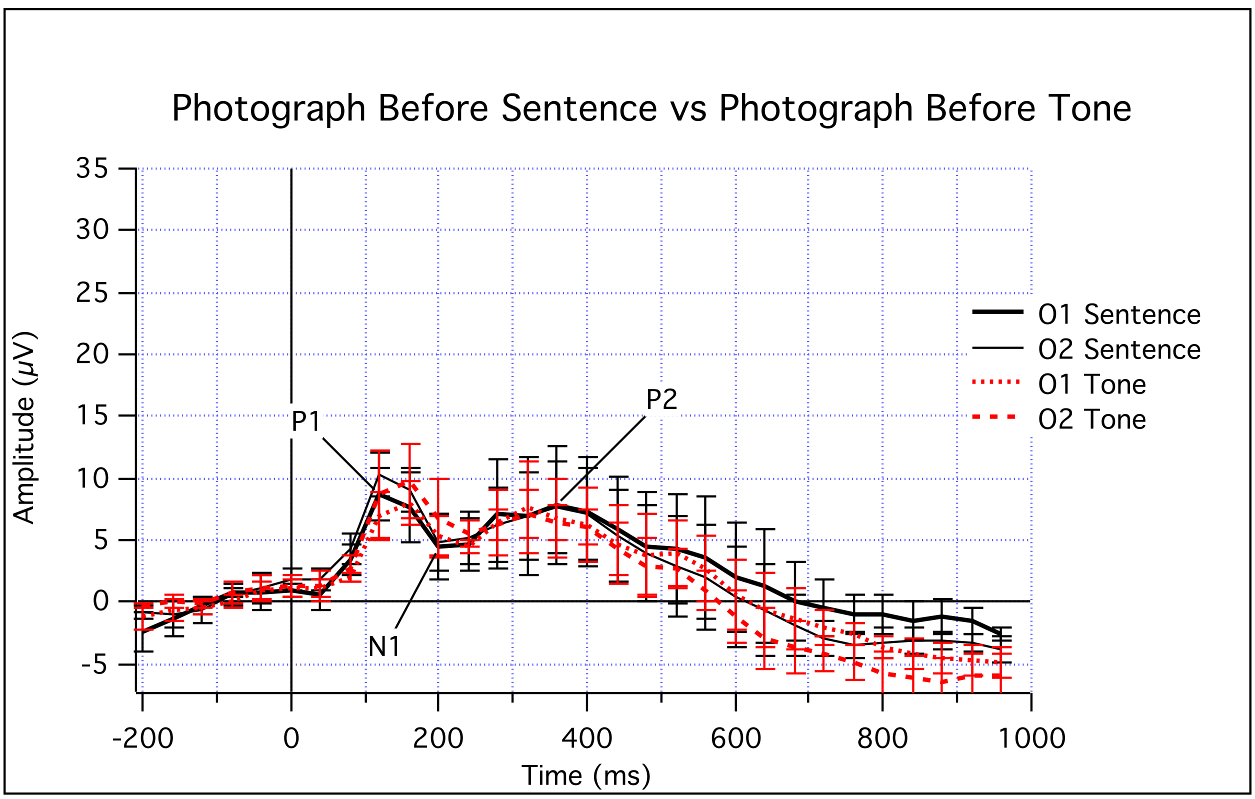

Figure 5; Averages for group of 3 participants who did Photograph/Tone and Photograph/Sentence conditions; Comparison from visual onset; O1(37), O2(40).

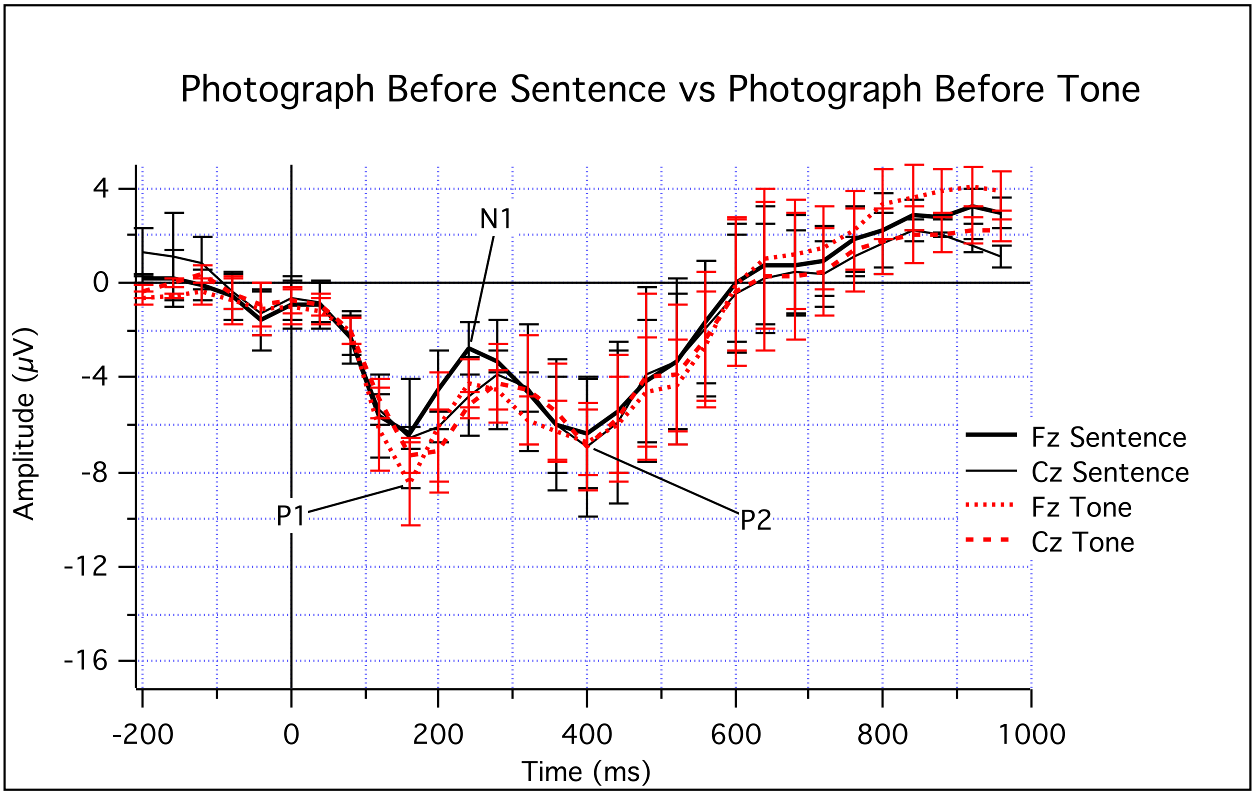

Figure 6; Averages for group of 3 participants who did Photograph/Tone and Photograph/Sentence conditions; Comparison from visual onset; Fz(4) Cz(65).

Sentence condition did not complete the Photograph /Tone condition as planned. It is possible future testing could bring out subtle differences in photograph processing when followed by linguistic as compared to non-linguistic auditory stimuli.

<u>Condition Comparisons from Auditory Onset</u>

A comparison of conditions from audio onset shows that photographs are contributing to brain responses in occipital and frontal sites during sentence presentation. Responses to sentences with photographs have a long slow negative wave at frontal sites with a corresponding positive inversion in occipital sites, which can be seen in both groups of participants (Figures 7, 8, 9 & 10). As with the previous graphs, error bars indicate standard error of the mean. For the group that also did the Fixation /Sentence

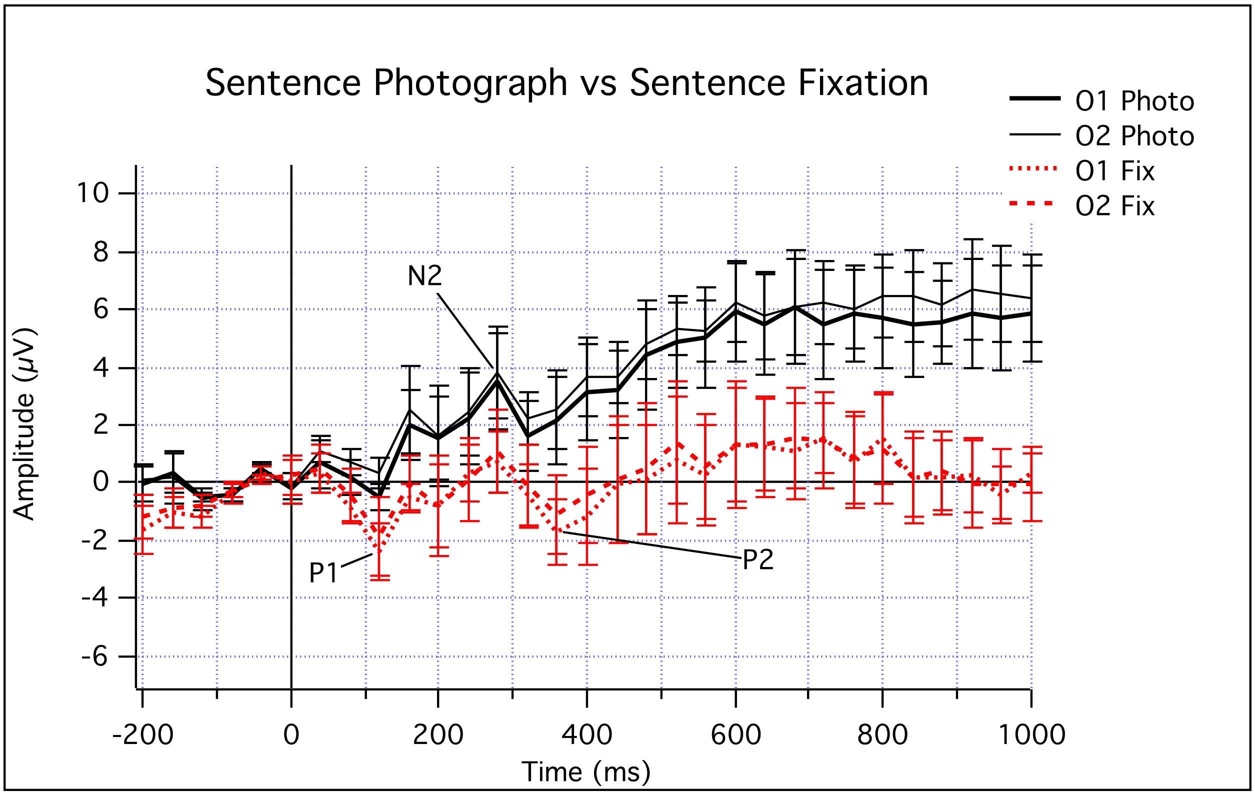

Figure 7; Averages for group of 4 participants who did Fixation /Sentence and Photograph /Sentence conditions; Comparison from auditory onset; O1(37), O2(40).

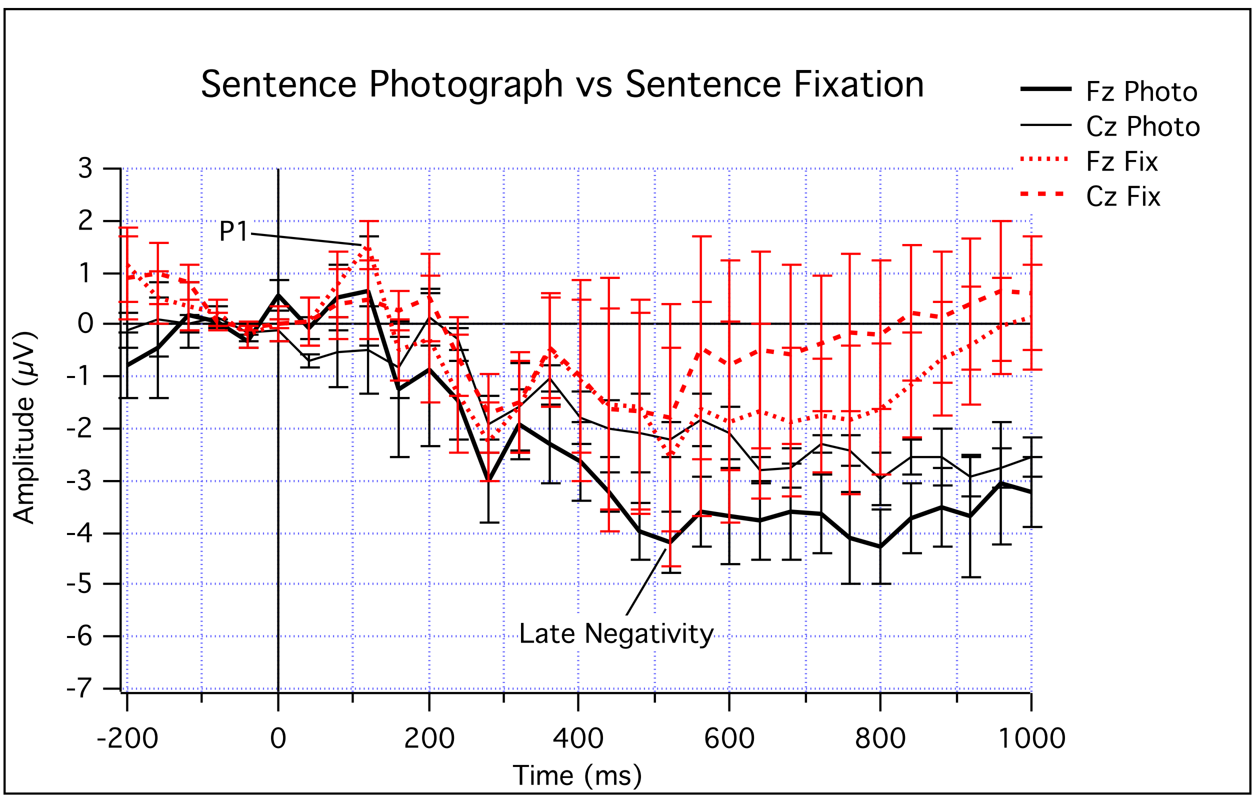

Figure 8; Averages for group of 4 participants who did Fixation/Sentence and Photograph/Sentence conditions; Comparison from auditory onset; Fz(4) Cz(65).

condition, this wave starts to diverge for the Photograph /Sentence condition between 300 ms to 400 ms from sentence onset at approximately 2 μV and rises to approximately 6 μV in occipital sites (Figures 7 & 8).

In comparing Photograph /Tone and Photograph /Sentence conditions from auditory onset, the responses for the Photograph /Sentence condition have been limited to include only the three participants who also did the Photograph /Tone condition. For the group that also did the Photograph /Tone condition, this late negative wave starts at approximately 360 ms after sentence onset in occipital sites and at 360 ms after sentence onset for Fz with little to no effect for Cz (Figure 9 & 10). The difference in group responses may be due to the fact that participants who did the Photograph /Tone condition were seeing the photographs for a second time around during the Photo /Sentence condition or possibly sample size is a factor. In addition, as predicted, the brain responses for sentences with related photographs differs from those of tones with photographs for this group (Figures 9 &10).

It is important to note that a comparison of averages from sentence onset for the four participant responses who did both the Photograph /Sentence and the Fixation / Sentence conditions show that this late wave is absent from the Fixation /Sentence condition in both occipital and frontal sites (Figures 7 & 8). Responses to sentences with a fixation are minimal in later time periods, hovering near zero μV within the standard error of the mean for points beyond P1 in occipital sites and for points beyond 350ms at frontal sites (Figure 7 & 8).

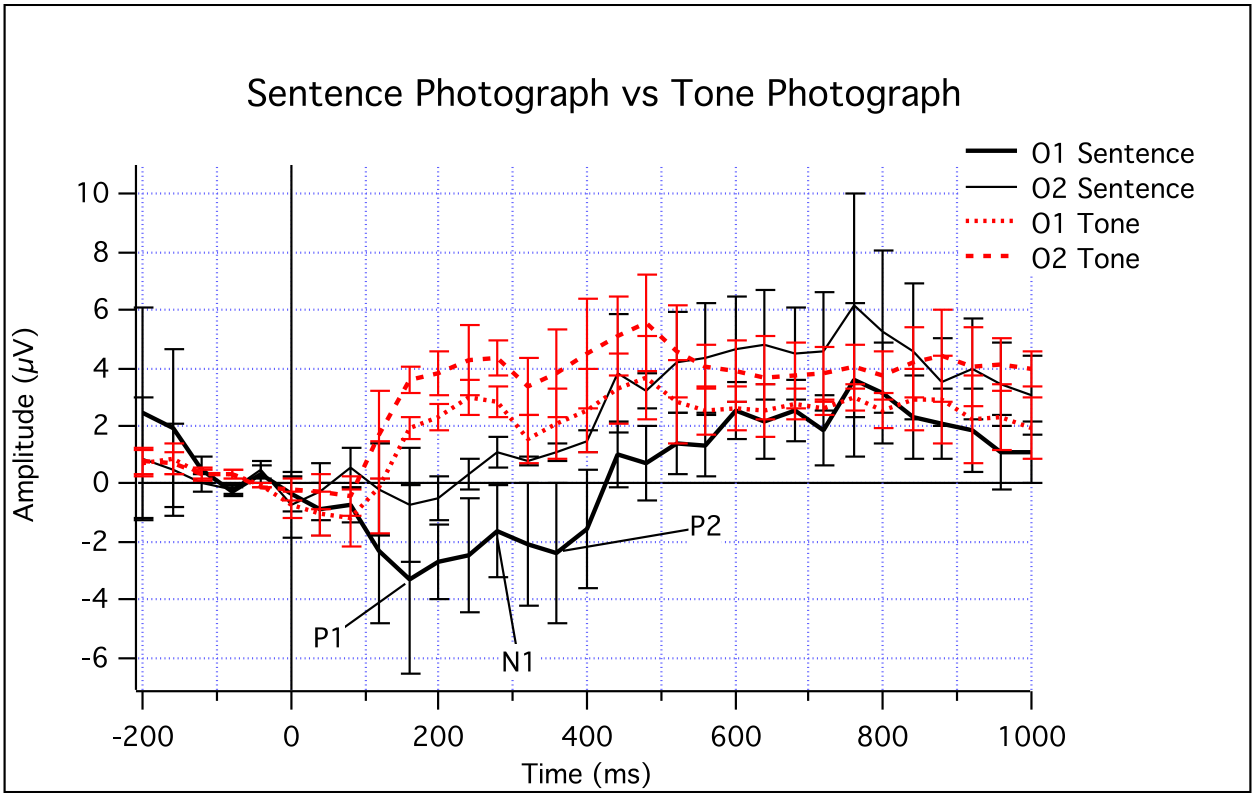

Figure 9; Averages for group of 3 participants who did Photograph/Tone and Photograph/Sentence conditions; Comparison from auditory onset; O1(37), O2(40).

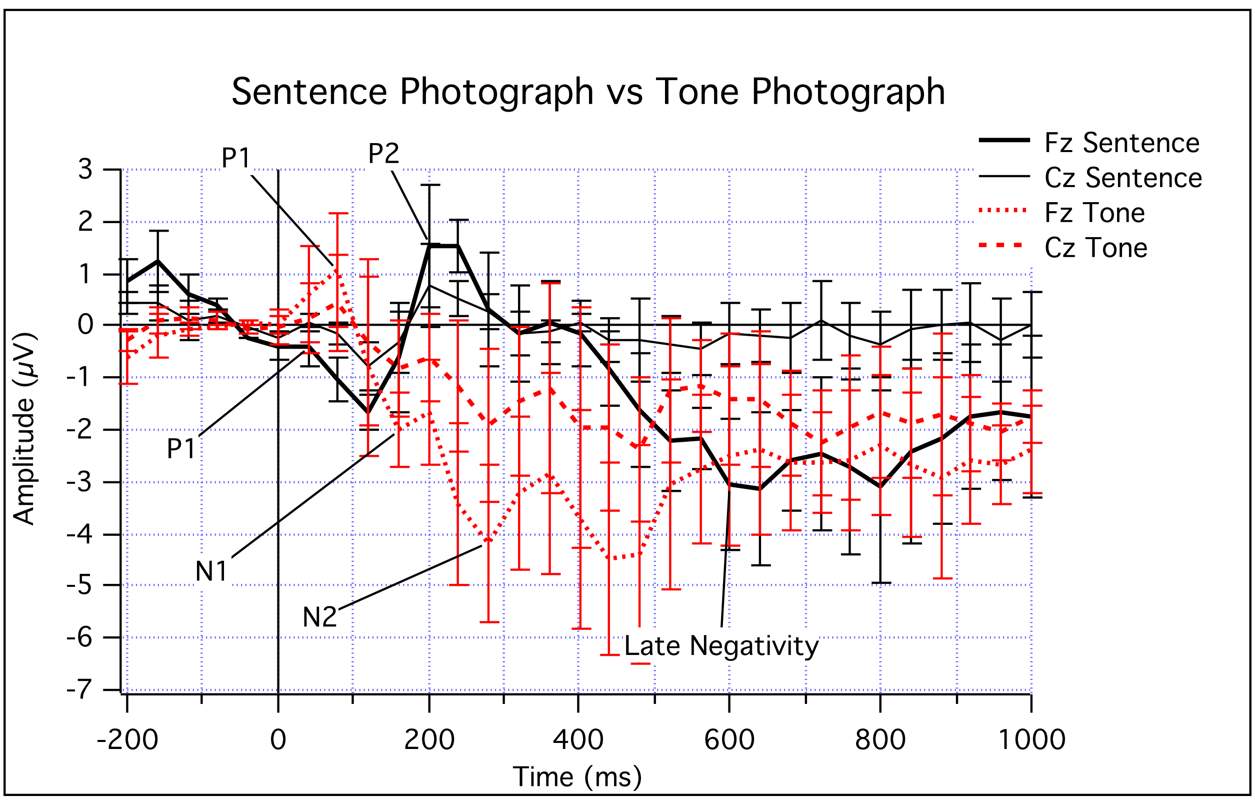

Figure 10; Averages for group of 3 participants who did Photograph/Tone and Photograph/Sentence conditions; Comparison from auditory onset; Fz(4) Cz(65).

Individual Participant Responses

Individual responses to the stimuli vary greatly among the eight children who participated in this study. Figures 11 and 12 show Individual responses in frontal and occipital sites from auditory onset. The small sample size makes generalizations tentative at this juncture, but there may be an age effect, which will be clarified with the addition of more participants. The two thirteen year old boys in this study have the smallest responses for both of the conditions in which they participated. The thirteen year old who did the fixation condition has overall responses within +/- 4 μV for the Sentence /Photograph condition and responses within +/- 3 μV for the Sentence /Fixation condition (Figure 11). The thirteen year old who did the Sentence /Photograph and Tone /Photograph conditions has overall responses within +/- 5 μV for both conditions (Figure 12). Whereas, the other children ranging in age from six to nine years old, all have larger amplitude responses, as large as +/- 12 μV for at least one of the conditions in which they participated.

When comparing the responses of participants who did the Sentence /Photograph and Tone /Photograph conditions (Figure 12) the only consistent difference in responses is approximately between 75 and 275 ms after audio onset (also Figure 9). In both these conditions the photograph is available to view for over 2000ms. In the Photoraph /Tone condition obligatory responses to the photographs occur well before the tone onset. Thus, evidence of processing these photographs (from occipital positivity) suggests voluntary attention to the photographs by some of the children.

The eight year old girl who participated in the Tone /Photograph condition (Figure 12) shows a larger occipital response to the photographs in the Tone /Photograph con-

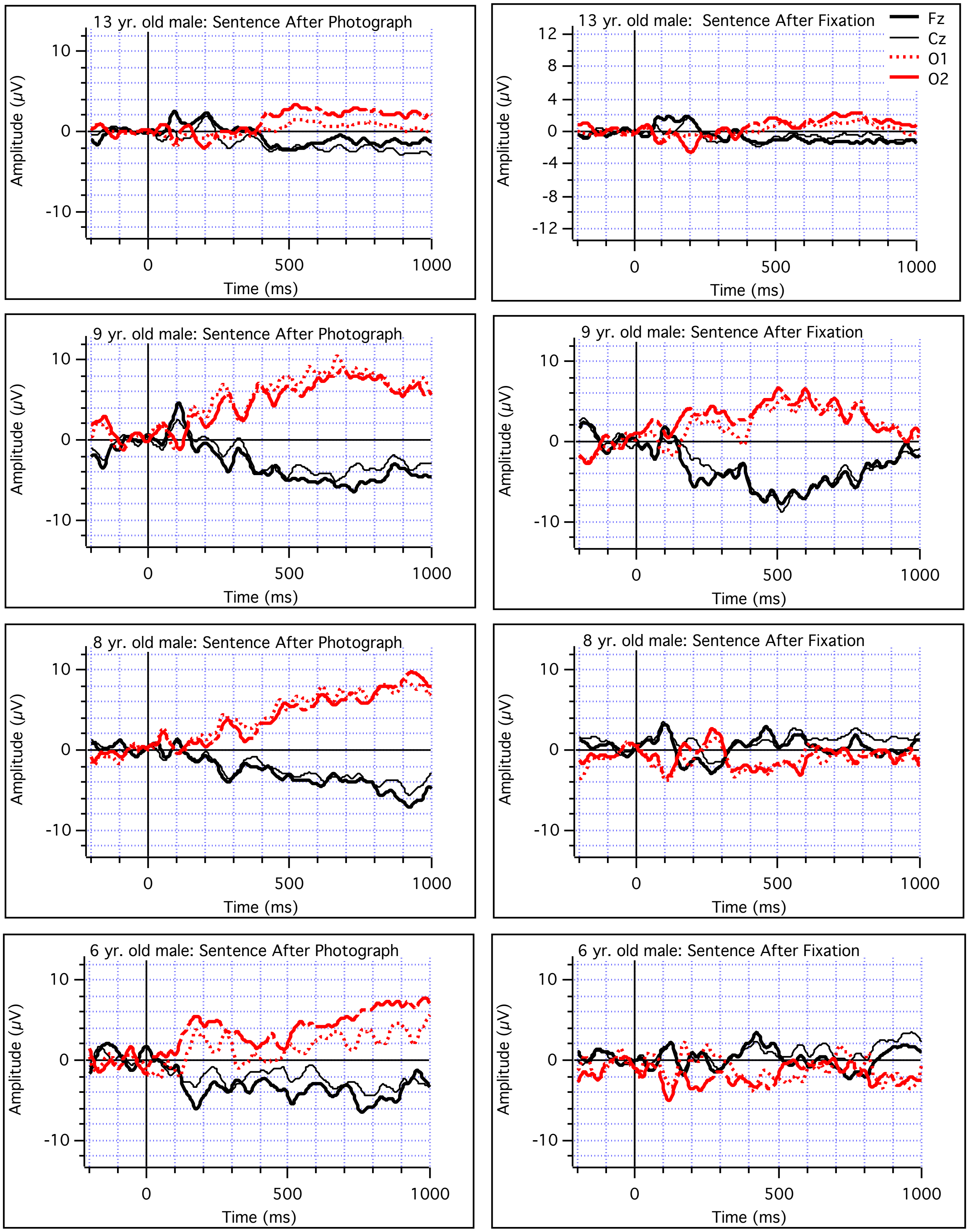

Figure 11. Individual responses from auditory onset for frontal and occipital sites. Group of children who did Sentence /Fixation condition, arranged from oldest to youngest.

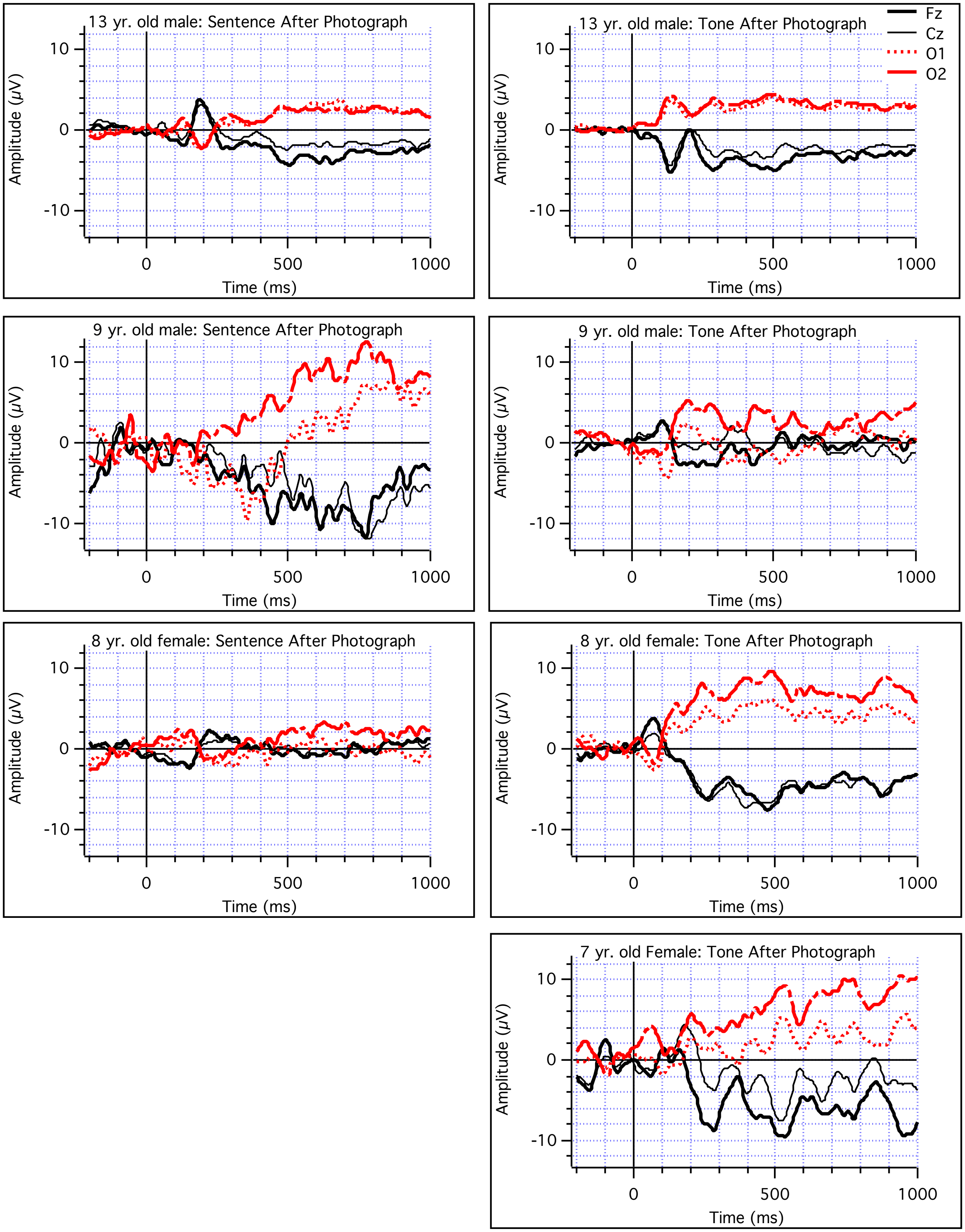

Figure 12. Individual responses from auditory onset for frontal and occipital sites. Group of children who did Tone /Photograph condition, arranged from oldest to youngest.

dition than in the Sentence /Photograph condition. It may be that this child is showing less audio visual integration, choosing to focus more on speech stimuli even in the presence of a related photograph and giving her attention to the photographs in the absence of any interesting auditory stimuli (leading to the large occipital response). However the other three participants, do appear to re-engage visual processing (as indexed by the occipital positivity) more strongly for the sentence following photograph condition as compared to the tone following photograph condition.

Discussion

This study shows variability in audio visual integration and in voluntary attention to visual stimuli for typical children, as shown by the results from auditory onset. Responses to sentences with related photographs vary among the individual children in obligatory auditory and visual responses. On average, sentences presented in the context of related photographs demonstrate that photographs are being re-assessed during sentence presentation, but not consistently across individuals. This suggests varying degrees of integration within the typical child population.

One variable appears to be the age of typical children in this study. Specifically, overall responses were smaller in amplitude for the older than the younger children. It is possible that the thirteen year olds were not as engaged generally due to the simplistic nature of the stimuli. However, adults often have smaller amplitude responses than children under 10 years of age (Shafer, Yu and Wagner; 2016). Thus, the smaller amplitude found for the teenage children may be related to maturation, such as more reliance on abstract thought than sensory input. It is currently unclear what has changed, but it is possible that maturation of cognitive systems leads to more specific engagement of cortex than for younger children.

Photographic processing during auditory processing is indicated in this study by a large positive slow wave over occipital sites with the inversion at frontal sites. Recall that the occipital positivity to the photograph onset falls back to baseline before the onset of the sentence or tone. The occurrence of the occipital wave during the presentation of tones which starts 1000 ms after photograph onset suggests that photographs are compelling images for children, producing spontaneous attention even after initial processing has been accomplished. Still, on average this wave is comparatively larger in the Sentence /Photograph condition, indicating that sentence presentation typically prompts a second look at the photograph to a greater extent than the tone. This pattern is expected considering that the sentences refer to the animals pictured in the photographs.

There is much less variability from unimodal visual onset. These results show that photographs produce a consistent double peaked response in occipital and frontal sites for typically developing children whether followed by a sentence or a tone, whereas the second peak is absent from responses to fixation symbols. This finding allows some speculation regarding the absence of a second peak found in Cantiani, et al., (2016) and in the previous syntax study motivating this study.

The study by Cantiani et al., (2016) suggests that some minimally verbal children with ASD are lacking in visual processing for photographs beyond early more automatic responses indexed by visual evoked obligatory responses. In addition speech that labeled the picture had little or no effect on the visual responses. This study presented the word labels 500 ms after the picture onset, and thus, it is difficult to evaluate whether differences in processing, particularly for the second peak between 400 and 800 ms after photograph onset, are related to processing the photograph alone or to integration of the photograph and speech. Similarly, the previous syntax study that motivated the present study presented sentences 300 ms after photograph onset.

There is research supporting both hypotheses for why children with ASD have diminished P2 responses to photographs. The fact that children with ASD are lacking adequate audio visual integration to meet requirements of visual processing that overlaps with speech processing is indicated by research done by Bebko (2006), De Gelder et al., (1991), Williams et al., (2004), Iarocci et al., (2010) and, Smith and Bennetto (2007). And, there is also research showing deficits for children with ASD in some aspects of visual processing alone, such as that done by Townsend et al., (2001).

Speed of processing may also have been a major contributing factor for attenuated P2 responses. The ERP study done by Cantiani et al., (2016) found later latencies along with intact amplitude for early components and diminished amplitude for later components for children with ASD. Later latencies of earlier components have been shown to correlate with smaller amplitudes of later components and to effect audio and visual pattern recognition with respect to emotions (Lerner et al.; 2013). Even if early components for sensory processing have similar amplitude to typical children, speed of audio and visual processing may affect higher order syntactic and semantic processing.

When children with ASD are tested in this paradigm, it is extremely likely that we will see the same slower latencies to photographs. But, it is possible that we may see more typical P2 amplitudes given that the children have 1000 ms before audio onset.

This would confirm that speed of processing photographs was the main obstacle for children on the autism spectrum. It would also suggest that children with ASD are capable of higher order cognition if the typical pace of realtime visual presentation is slowed down considerably.

If children with ASD do not have P2 responses before audio onset, we can assume they will not show any voluntary attention to photographs at later time periods, since fixation symbols which lacked a P2 response from visual onset also lacked the late slow positivity in occipital sites during sentence presentation. If children with ASD do show a P2 response from photograph onset and also have evidence of voluntary attention to photographs from auditory onset, we will be able to determine if sentences are prompting a re-assessment of photographs, since the late slow positive wave is typically larger in the Photograph /Sentence condition. But, this result would not be expected given what we know of the sensory integration difficulties of children with ASD.

Typical language acquisition requires appropriate attention, perception and memory for pattern recognition and appropriate processing speed and synchronized timing across sensory modalities These basic functions are in a feedback loop with higher levels of cognition that are also needed to acquire all aspects of language, such as the bridging theories needed to understand sarcasm. Presumably native neurology aids language acquisition at every sequential stage of development. Intense World Syndrome Theory (Markram, RInaldi and Markram, 2007) and other proposals (e.g., Entrainment deficits, Jochaut, et al., 2015) are formulated to account for ASD. However, the neural underpinnings of cross-modal perceptual experience in children with typical development is largely unknown. Thus, it is difficult to understand where processing breaks down in ASD without having a better understanding of typical development. The current study provided some information regarding obligatory audio visual responses as well as voluntary visual attention to photographs related to sentences.

Limitations and Future Directions

This study could be altered, in order to confirm that the difference found for sentence processing conditions in typical children results from the integration of photograph content as opposed to differences in the qualities of the image available to view during sentence presentation. To this end, a degraded image that retains the qualities of the photographs without discernible forms could be a useful replacement for fixation symbols.

Testing for this study is ongoing. At this point, it is unclear to what extent the small sample size and wide range of ages may be affecting results. Further testing of more participants will clarify these findings.

In addition, future plans for this study include testing children with ASD who exhibit a range of verbal ability. When children with ASD are tested, it will become clear if the differences in their brain responses for the original paradigm are due to differences in photograph processing or sentence processing or the timing of the overlap between the two. Comparing sentences with and without photographs will show if children with ASD integrate photographs during sentence processing. And, responses to sentences followed by photographs after 1000 ms can be compared to the 300 ms paradigm to see if speed of processing was an obstacle for the children with ASD who participated in the original study.

Additional analysis of the data would also be helpful. It would be interesting to discern if children with ASD have typical entrainment for unimodal conditions versus multimodal conditions. In the future, an analysis of alpha, beta, gamma, delta and theta oscillations of the EEG could provide information on sensory linguistic entrainment. In addition, audiovisual coordination could be ascertained by analyzing eye blink data captured by EEG with respect to prosodic pauses.

Future studies would benefit from a sensory profile for participants with ASD, in order to obtain more detailed information on how degrees of sensory processing differences correlate with degrees of language differences. Donahue et. al., (2012) have shown that the degree of sensory integration differences correlate with the degree of self reported ASD symptoms. More needs to be known about how atypical sensory integration affects language acquisition throughout the process of interpreting speech events, and being able to respond in conversation. In addition, more research is needed on the consequences of auditory dominance aover visual dominance for multi sensory pattern recognition in terms of acquiring pragmatic cues, semantic references and phoneme recognition.

Future studies investigating the connection between additional physiological sensory abilities may also shed light on the lack of pragmatic use for individuals with ASD. In phonology the ability to produce phonemes effects the ability to perceive phonemes. Poor proprioception and motor control may be effecting the ability of individuals with

ASD to perceive facial expressions and gestures needed for pragmatic use and acquisition.

If a core deficit is to be found for autism, it would need to restrict the use and acquisition of pragmatics, and in certain cases interfere with the typical course of phonological, syntactic and semantic acquisition as well. Research has shown that some of the earliest foundations of language acquisition, defined here as the process of interpreting and using speech events, may be traced back to the first half year of life to native mechanisms of entrainment and sensory binding. Entrainment and sensory binding are atypical throughout the lifetime of individuals with ASD and may start them off from the earliest stages on a divergent path of language acquisition that does not become obvious behaviorally until later points in life.

Conclusion

The rich image such as these photographs does not produce the same type of attenuation in ERP response for typical children as the simple fixation symbol. An examination of ERPs at frontal and occipital sites to photograph onset resulted in a consistent double peaked response whether or not photographs were followed by sentences or tones. Responses to fixation symbols had a P1 but lacked a P2 response.

Brain responses to audio onset had less consistent results across participants, but generally showed a late slow positive wave at occipital sites in conditions where photographs were available to view. On average larger responses were found to sentence onset than to tone onset. The average difference in brain responses for sentences with photographs compared to sentences with fixation suggests that a sentence presented in the context of a related photograph is not typically processed as the same kind of event as a sentence presented with a fixation symbol.

This pattern of occipital positivity can be used to infer what children are doing with their attention in cross-modal passive tasks. The variability in responses across participants during auditory presentation suggests that visual attention to previously presented visual stimuli is under the participants volitional control, rather than being automatic. But, typical children will re-assess available visual stimuli well after the completion of initial processing when prompted by related speech stimuli. These findings suggests that typical children's brain responses show an effect of voluntary attention to photographs, which is enhanced when prompted by the introduction of a related sentence. The more we understand about intermodal linguistic processing for typically developing children, the more we will be able to discover what has gone awry in language acquisition for children on the autism spectrum.

Bibliography

American Psychiatric Association [APA]. (2013). *Diagnostic and Statistical Manual of Mental Disorders*, 5th Edn. (Arlington: American Psychiatric Association;).

Asht, S. and Dass, R. (2012). *Pattern Recognition Techniques: A Review*. International Journal of Computer Science and Telecommunications, Volume 3, Issue 8, August

Ayres, A.J. (1972). *Improving Academic Scores Through Sensory Integration.,* Journal of Learning Disabilities 5 (6): 338–343.

Baren-Cohen, S.; Leslie, A.M.; Frith, U. (1985). Does the autistic child have a "theory of mind"?. Cognition V.21, Issue 1, October P 37-46

Barnes, R.; Jones, M. R. (2000). *Expectancy, Attention, and Time*. Cognitive Psychology, V 41, Issue 3, Nov., P. 254-311

Bebko, J.M.; Weiss J.A.; Demark, J.L.; Gomez, P. (2006). *Discrimination of temporal synchrony in intermodal events by children with autism and children with developmental disabilities without autism.* J Child Psychol Psychiatry.Jan; 47(1):88-98.

Bogdashina, O. (2013). *Sensory theory in autism makes sense: A brief review of the past and present research*. OA Autism 2013 Mar 01;1(1):3

Calvert, G.; Spence, C.; Stein, B.E. (2004). *The Handbook of Multisensory Processes,* MIT Press.

Cantiani, C; Choudhury, N.A.; Yan H. Yu, Y.H.; Shafer, V.L.; Schwartz, R.G.;Benasich, A.A.(2016). *From Sensory Perception to Lexical-Semantic Processing: An ERP Study in Non-Verbal Children with Autism.* Plos One August 25,

Ciesielski, K.T.; Courchesne, E.; Elmasian, R. (1990). *Effects of focused selective attention tasks on event-related potentials in autistic and normal individuals*. Electroencephalography and Clinical Neurophysiology. Volume 75, issue 3, March, Pages 207–220.

Clark, H.H.; Gerrig R.J. (1984) *On the pretense theory of irony* . Journal of Experimental Psychology: General 113. 1 (Mar): 121-126

Courchesne, E.; Lincoln, A.J.; Kilman, B.A.; Galambos, R. (1985b). *Event-related brain potential correlates of the process- ing of novel visual and auditory information in autism*. Journal of Autism and Developmental Disorders, 15, 55–76.

De Gelder, B., Vroomen, J., Van Der Heide, L. (1991). *Face regognition and lip-reading in autism.* J. Autism and Developmental Disorders 18, 493-504

Donohue, S.E.; Darling, E.F.; Mitroff, S.R. (2012). *Links between multisensory processing and autism*.

Eigsti, I.; de Marchena, A.B.; Schuh, J.M.; Kelley, E. (2011). *Language acquisition in autism spectrum disorders: A developmental review.*, Research in Autism Spectrum Disorders 5, 681–691.

Eigsti, I.; Bennetto. L. (2009). *Grammaticality judgments in autism: Deviance or delay*. Journal of Child Language/ Volume 36 / Issue 05 / November, pp 999-1021.

Ferri, R.; Elia, M.; Agarwal, N.; Lanuzza, B.; Musumeci, S.A.; Pennisi, G. (2003). *The mismatch negativity and the P3a compo- nents of the auditory event-related potentials in autistic low- functioning subjects*. Clinical Neurophysiology, 114, 1671–1680.

Foss-Feig, J.H.; Kwakye, L.D.; Cascio, C.J.; Burnette, C.P.; Kadivar, H.; Stone, W.L.; Wallace, M.T. (2010). *An extended multisensory temporal binding window in autism spectrum disorders*. Exp Brain Res 203(2):381–389.

Frith, U.; Happe, F.; (1994). *Autism: beyond "theory of mind"*. Cognition V 50, Issues 1-3 April-June Pages 115-132.

Gomot, M.; Giard, M.; Adrien, J.; Barthelemy, C.; Bruneau, N. (2002). *Hypersensitivity to acoustic change in children with autism: Electrophysiological evidence of left frontal cortex dysfunctioning.* Psychophysiology, 39, 577–584.

Grice, H.P. (1975) *Logic and conversation*. In Cole, P., and J.L. Morgan, eds. Speech Acts. New York: Academic Press, 41–58.

Haith, M.M.; Hazan, C.; Goodman, G.S. (1988). *Expectation and Anticipation of Dynamic Visual Events by 3.5-Month-Old Babies*. Child Development, Vol. 59, No. 2 (Apr), pp. 467-479.

Hecht, D.; & Reiner, M. (2009). *Sensory dominance in combinations of audio, visual and haptic stimuli.* Experimental Brain Research, 193 (2), 307–314

Iarocci, G.; Rombough, A.; Yager, J.; Weeks, D. J;, & Chua, R. (2010). *Visual influences on speech perception in children with autism.* Autism, 14(4), 305-320.

Jones, M.R. (1976). *Time, our lost dimension: Toward a new theory of perception, attention, and memory.* Psychological Review, 83, 323–355.

Jorgensen, J.; Miller, G.A.; Sperber, D. (1984). *Test of the mention theory of irony*. Journal of Experimental Psychology: General 113. 1 (Mar): 112-120.

Jochaut, D.; Lehongre, K.; Saitovich, A.; Devauchelle, AD.; Olasagasti, I.; Chabane, N.; Zibovicius, M.; Giraudi, AL. (2015). *Atypical coordination of cortical oscillations in re-*

sponse to speech in autism. Front Hum Neurosci. 2015; 9: 171. Published online Mar 27.

Kanner, L. (1949). *Problems of Nosology and Psychodynamics of Early Infantile Autism.* American Journal of Orthopsychiatry 19: 416–26.

Kuhl. P.K, (2004), *Early language acquisition: cracking the speech code*. Nat Rev Neurosci. Nov; 5(11):831-43.

Lepisto, T.; Kajander, M.; Vanhala, R.; Alku, P.; Houtilainen, M.; Naatanen, R.; Kujala, T. (2008). *The perception of invariant speech features in children with autism.* Biol Psychol., Jan;77(1):25-31.

Lerner, M.D.; McPartland, J.C.; Morris, J.P. (2013)*. Multimodal emotion processing in autism spectrum disorders: An event-related potential study.* Developmental Cognitive Neuroscience 3, 11–21.

Lewkowicz, D.J.; Hansen-Tift, A.M. (2012). *Infants deploy selective attention to the mouth of a talking face when learning speech.* Proc Natl Acad Sci U S A., Jan 31;109(5):1431-6.

Liberman, A.M; Mattingly, I.G (1985). *The motor theory of speech perception revised*. Cognition **21**(1):1–36.

Marco, E.J.; Barett, L.; Hinkley, N.; Hill, S.S.; Nagarajan, S.S. (2011). *Sensory Processing in Autism: A Review of Neurophysiologic Findings.* Pediatr Res. May;69(5 Pt 2):48R-54R.

Markram, H.; Rinaldi, T.; Markram, K. (2007). *The Intense World Syndrome—an alternative hypothesis for autism.* Front. Neurosci, Oct 15.

Martineau, J.; Garreau, B.; Barthelemy, C.; Lelord, G. (1984). *Evoked potentials and P300 during sensory conditioning in autistic children.* Annals of the New York Academy of Sciences, 425, 362–369.

McAuley, J. D.; Jones, M. R.; Holub, S.; Johnston, H. M.; Miller, N. S. (2006). *The time of our lives: Life span development of timing and event tracking*. Journal of Experimental Psychology: General, Vol 135(3), Aug, 348-367.

Nakano, T.; Kitazawa, S. (2010) *Eyeblink entrainment at breakpoints of speech.* Experimental Brain Research. September, Volume 205, issue 4 pp 577-581.

Nakano, T.; Kato, N; Kitazawa, S. (2011). *Lack of eyeballing entrainments in autism spectrum disorders*. Neuropsychologia. 06.007. Epub Jun 12.

Novick, B.; Kurtzberg, D.; Vaughn, H.G., Jr. (1979). *An electrophysiologic indication of defective information storage in childhood autism.* Psychiatry Research, 1, 101–108.

Novick, B.; Vaughan, H.G.; Jr., Kurtzberg, D.; Simson, R. (1980). *An electrophysiologic indication of auditory processing defects in autism*. Psychiatry Research, 3, 107–114.

Oades, R.D.; Walker, M.K.; Gaffe, L.B.; Stern, L.M. (1988). *Event-related potentials in autistic and healthy children on an auditory choice reaction time task*. International Journal of Psychophysiology, 6, 25–37.

Oconnor, K.; Hamm, J.P.; Kirk, I.J. (2006). *Neurophysiological responses to face, facial regions and objects in adults with Asperger's syndrome: An ERP investigation*, ijosycho. 12.001.

Rhoshel, K.; Lenroot, P.K.Y. (2013). *Heterogeneity within autism spectrum disorders: what have we learned from neuroimaging studies?* Front. Hum. Neurosci., 30 October.

Shafer, V.; Yu, Y.H; Wagner, M. (2015). *Maturation of cortical auditory evoked potentials (CAEPs) to speech recorded from frontocentral and temporal sites: three months to eight years of age.* Int J Psychophysiology, Feb, 95(2): 77-93.

Smith E.G.; Bennetto, L.; (2007). *Audiovisual speech integration and lipreading in autism.* Journal of Child Psychology and Psychiatry., 48:813–821

Sun, L.; Grützner, C.; Bölte,S.; Wibral, M.; Tozman, T.; Schlitt, S.; Poustka, F.; Singer, W.; Freitag,C.M.; Uhlhaas, P.J. (2012). *Impaired Gamma-Band Activity during Perceptual Organization in Adults with Autism Spectrum Disorders: Evidence for Dysfunctional Network Activity in Frontal-Posterior Cortices,* The Journal of Neuroscience, 11 July, 32(28).

Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.

Townsend, J.; Westerfield, M.; Leaver, E.; Makeig, S.; Jung, T.P.; Pierce, K,; Courchesne, E. (2001). Event-related brain response abnormalities in autism: evidence for impaired cerebello-frontal spatial attention networks. Cognitive Brain Research 11 127–145.

Wang, J.; Barstein, J.; Ethridge, L.E.; Mosconi, M.W.; Takarae, Y. Sweeney, J.A. (2013). *Resting state EEG abnormalities in autism spectrum disorders.* Journal of Neurodevelopmental Disorders, 5:24.

Whalen, D.H., Levitt, A.G., Wang, Q. (1991). I*ntonational differences between the reduplicative babbling of French and English learning infants.* Journal of Child Language, Volume 18, Issue 3, October, pp. 501-516

Williams, J.H.G.; Massaro, D.W.; Peel, N.J.; Bosseler, A. Suddendorf, T. (2004). *Visual–auditory integration during speech imitation in autism. Research in Developmental Disabilities, V. 25, Issue 6, Nov.-Dec., P. 559-575*