

# A neural model of sensory interactions in young neurotypical and ASD children

M. Monti<sup>1</sup>, S. Molholm<sup>2</sup>, and C. Cuppini<sup>1</sup>

<sup>1</sup> Department of Electrical, Electronic, and Information Engineering Guglielmo Marconi, University of Bologna, Italy

<sup>2</sup> The Cognitive Neurophysiology Laboratory, Albert Einstein College of Medicine, New York, USA

**Abstract**—The brain at birth is characterized by a default state of competition among the senses [1], and integrative abilities are acquired only later in life, based on experience. Recently, Crosse and colleagues [2] in a bisensory reaction time task, with auditory (A) and visual (V) stimuli, presented alone or together (AV), in a random sequence, showed that children’s reaction times (RTs) to congruent AV stimuli did not differ significantly from unisensory RTs, but differences emerged between ASD and neurotypical (NT) subjects. In this work, we implemented a neurocomputational model to simulate the RTs and analyse the behavioural responses of ASD children (6-7 years of age) and their NT counterpart. The model suggests that the comparable RTs found in unisensory and multisensory conditions could be interpreted by the default competition among the senses; this default state can be implemented via mutual inhibition between primary sensory areas; and the shift toward the classical multisensory facilitation, observed in adults, is the result of the inhibitory cross-modal connections becoming excitatory after an extended multisensory experience. Moreover, model results suggest that differences between ASD and NT children are due to a stronger cross-modal inhibition in 6-7 year-old ASD children. Finally, this neurocomputational model allowed investigation of the temporal profile of interactions among stimuli of different sensory modalities. These findings link the perceptual framework delineated by several empirical results to a plausible neural implementation.

**Keywords**—Neural Modelling, Autism, Sensory Interactions.

## I. INTRODUCTION

MULTISENSORY integration (MSI) is the mechanism exploited by the brain to deal with information from different sensory systems, and optimizing its responses in case of redundant multimodal cues compared to the individual unisensory components, when presented separately. The principles ruling MSI in the spatial domain, like the inverse effectiveness and the multisensory enhancement in case of congruent multisensory stimuli, have been known for a long time [3], [4]. In the temporal domain great effort has been dedicated to identify the conditions for a multisensory facilitation (the so-called temporal binding window, TBW) [5], [6], but only recently the competition among stimuli incongruent in time has been analysed: Cuppini et al. [7] addressed MSI in the temporal domain for healthy adult subjects. Additionally, MSI is not innate: the brain at birth is characterized by a default state of competition among the senses, and integrative abilities are acquired only later in life, based on sensory experience [1].

In the last decade, numerous studies dealing with MSI focused on its alterations in subjects with neurodevelopmental disorders. Among these, individuals with Autism Spectrum Disorders (ASD) have shown a lack of benefit from multiple sources of information as their neurotypical (NT) peers do [8]–[12]. These deficits have been observed in case of simple AV

detection task, involving low-level stimuli [2], [13], but also with audiovisual (AV) speech stimuli, thus suggesting that perceptual deficits may be responsible for higher order core symptoms characterizing ASD [12].

Recently, Crosse and colleagues [2] implemented a bisensory reaction time task to characterize the MSI maturation, in NT and ASD, from childhood to the adult age, and analyse the switch cost (the difference between RTs in case of stimuli preceded by a different sensory modality, switch trials, and RTs in repeat trials, when the preceding stimuli are of the same modality). In their experimental setup, auditory (A) and visual (V) stimuli were presented in random sequence, alone or together (AV), and subjects were required to respond (pressing a button) as soon as they perceived a stimulus. They showed that children’s reaction times (RTs) to congruent AV stimuli did not differ significantly from unisensory RTs. Similarly to [1], also these authors suggested that this behaviour could be ascribed to a default competitive interaction among sensory modalities. Postnatal multisensory experience is thought to craft cross-sensory projections between sensory cortices, promoting the maturation of integrative abilities and shift the competition to facilitative cross-modal interactions. Authors highlighted that this shift is delayed in ASD children.

Despite the numerous developmental studies, reporting reduced multisensory abilities in younger children, there is little evidence of such competitive cross-modal architecture. Moreover, to our knowledge, there is no study investigating the mechanisms underlying MSI in the temporal domain for children and/or for any pathological condition. Therefore, here, we used a computational modelling framework to reproduce the data of Crosse et al. [2], investigating the nature of cross-modal interactions in the youngest subjects (6-7 years of age), characterizing the differences between NT and ASD children, and the mechanisms underlying MSI in the temporal domain.

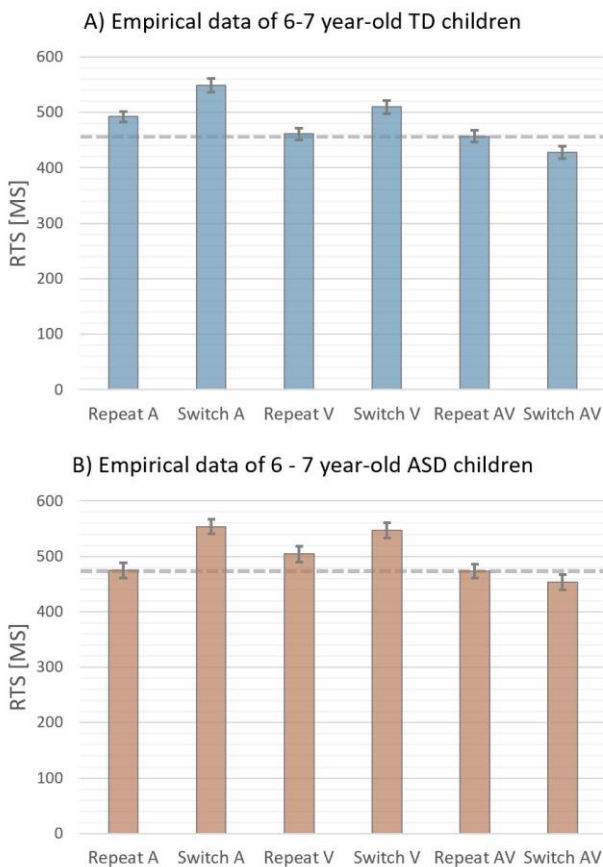
### A. Data analysis

To identify the neural mechanisms to be implemented in the network, first we analysed the empirical results of [2], based on the age group, in the case of unisensory and multisensory stimulations.

As it can be seen in Fig. 1, data of 6-7 year-old subjects do not show a multisensory behavioural facilitation. Indeed, responses (RTs) were not significantly faster in the multisensory condition compared to the unisensory ones: in NT children (Fig. 1A), RTs in AV repeat trials were comparable with RTs in the visual repeat condition; in ASD children (Fig. 1B), RTs in AV repeat trials were comparable with RTs in the auditory repeat condition. Moreover, RTs in case of multisensory switch trials, do not differ neither from

AV repeat, nor from unisensory repeat condition. This pattern of behaviour not only denotes the absence of a multisensory facilitation, but it suggests an active competition between the two sensory modalities: the fact that behavioural responses in the multisensory repeat condition are comparable with RTs in the fastest unisensory repeat condition, suggests that in case of AV stimulation, the two sensory channels compete, and only one of them can drive the response and determine the RTs. In agreement with [1], [2], these results strongly suggest the existence of an initial cross-modal competition between the sensory modalities. During the development, this competition should be superseded by the classical multisensory facilitation, as the result of the multisensory experience a subject was exposed to.

In our model, this default cross-modal competition was implemented via inhibitory cross-modal projections, as shown in Fig. 2. These synapses produce a fast competition between sensory modalities, and, balancing the integration occurring in the M region (due to the excitatory feedforward connections projecting on it), nullify any multisensory facilitation.



**Fig. 1:** Analysis of unisensory and multisensory RTs for both NT and ASD children of 6-7 years of age. A) In NT children of 6-7 years of age, RTs in the multisensory repeat trials were comparable with RTs in the visual repeat condition. B) In ASD children of 6-7 years of age, RTs in the multisensory repeat trials were comparable with RTs in the auditory repeat condition. It is evident that multisensory RTs were not faster than the faster of the unisensory conditions, but exactly comparable to it. This behavioural data could only be explained by a competition between the two sensory modalities.

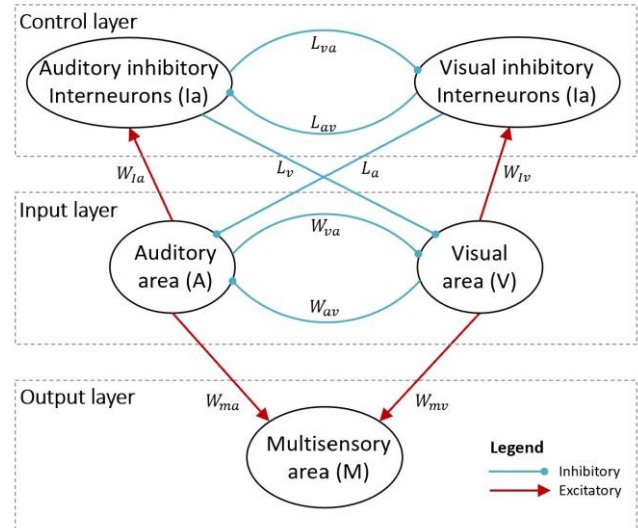
## II. METHODS

### A. Model architecture

The model has been realized to simulate the behavioural task implemented by [2], which focuses on multisensory

interactions in the temporal domain: subjects (either adults or children, NT or ASD) were required to respond to auditory and visual stimuli, alone or combined, presented in a random sequence, with an inter-stimulus-interval (ISI) varying between 1 s and 3 s. Given the experimental set-up in [2], we do not require either multiple units sensitive to different spatial positions in each sensory area, nor do we need multiple sensory regions sensitive to different input features.

The model we developed is represented in Fig. 2. The network is composed of three layers.



**Fig. 2:** Model structure. A and V represent the auditory and visual regions, responsible for the first sensory processing, and implementing the input layer. They exchange direct inhibitory synapses ( $W_{va}$  and  $W_{av}$ ), implementing a cross-modal competition. M is the multisensory output region, excited by the input layer, through feedforward connections ( $W_{ma}$  and  $W_{mv}$ ). It is responsible for generating the simulated reaction times (RTs) to the external stimuli.  $I_a$  and  $I_v$  are unimodal inhibitory areas, excited by the input layer (through  $W_{Ia}$  and  $W_{Iv}$ ) and implementing a control mechanism via feedback inhibitory synapses ( $L_a$  and  $L_v$ ) and a winner-takes-all mechanism ( $L_{va}$  and  $L_{av}$ ).

A and V represent the auditory and visual regions, respectively, responsible for the first sensory processing of external stimuli, and implementing the “input layer”. A and V exchange reciprocal inhibitory synapses,  $W_{va}$  and  $W_{av}$ , with a fast dynamic, implementing the default cross-modal competition.

M is the multisensory output region (“output layer”), responsible for generating the simulated reaction times (RTs) to the external stimuli. M, which receives the converging excitatory projections ( $W_{ma}$  and  $W_{mv}$ ) from the two areas in the input layer, may correspond to multisensory associative cortices, including, for example, the Posterior Parietal Cortex (PPC).

Ia and Iv are unisensory inhibitory areas representing a “control layer”. They are excited by the input layer (through  $W_{Ia}$  and  $W_{Iv}$ ) and implement a cross-sensory competition between the two sensory modalities, characterized by a slow dynamics: if stimulated by an input of the corresponding sensory modality, these modality-specific inhibitory interneurons produce a long lasting inhibition of the input region processing external stimuli of the other sensory modality, via the inhibitory feedback projections ( $L_a$  and  $L_v$ ), and the winner-takes-all mechanism ( $L_{va}$  and  $L_{av}$ ). It is worth noting that, since the dynamics of this mechanism is slow, the processing of the current stimulus is affected by the previous

one, when of the opposite sensory modality. This competitive layer can be implemented through higher order regions, for example located in the medial Prefrontal Cortex (mPFC) or the Posterior Cingulate Cortex (PCC).

### B. Assessment of network performance

First, as described above, to discriminate between alternative architectures for the network, we performed an analysis of the data collected by [2], based on the age group, in the case of unisensory and multisensory stimulation.

Once we identified the cross-modal inhibition as the more plausible neural architecture to explain these behavioural data, we performed several simulations to test the network behaviour. To this end, we presented sequences of unisensory (A or V) and congruent AV stimuli, in randomized order. The onset, duration, and presentation rate (ISI) of the stimuli were chosen to mimic the experimental setup of [2], whose data are compared to our results. Specifically, external stimuli are excitatory inputs with a duration of 60ms, presented at a rate of every 2s (the mean ISI in [2]). The activity elicited in the output region was compared with a fixed threshold,  $\varphi$  (10% of the maximum neurons' activity), to evaluate the simulated RTs, computed as the time interval between the instant of input presentation, and the instant when the evoked activity in the output area reaches the threshold. These results were analysed separately based on the respective input modality.

## III. RESULTS

In the following we present our results, critically analysing the model behaviour. This allows us to understand (i) what mechanisms are responsible for behavioural data in children and (ii) how such mechanisms differ in ASD.

TABLE I: SYNAPTIC WEIGHTS

NT children			
$L = 3$	$L_{a_{NT}} = 0.09$	$W_{av_{NT}} = -0.45$	$W_{ma_{NT}} = 1$
$WI = 1.4$	$L_{v_{NT}} = 0.07$	$W_{va_{NT}} = -0.45$	$W_{mv_{NT}} = 1.72$
ASD children			
$L = 3$	$L_{a_{ASD}} = 0.13$	$W_{av_{ASD}} = -0.8$	$W_{ma_{ASD}} = 1.1$
$WI = 1.4$	$L_{v_{ASD}} = 0.05$	$W_{va_{ASD}} = -0.8$	$W_{mv_{ASD}} = 1.3$

### A. Testing model simulations against empirical data

Once we identified the cross-modal inhibition as the main mechanism to be implemented, the next step was to determine the exact parameters' values that allow to reproduce RTs reported in [2], for both ASD and NT children at 6-7 years of age. To this aim, several sets of parameters were tested, and the simulated RTs were compared with the experimental median RTs extracted from the population of real subjects.

Fig. 3 illustrates the comparison between simulated and real RTs, obtained with the best set of parameters for all the six input configurations, for both ASD and NT children. It is worth noting that, among all the parameters, only the synaptic weights differ between the two groups, and are reported in Table I.

In particular, model results suggest that differences between ASD and NT children can be explained by a stronger cross-modal inhibition in 6-7 year-old ASD children, as can be noted from the more negative value assumed by  $W_{va}$  and  $W_{av}$  in the ASD simulations (Table I). We made this assumption based on the delayed maturational trajectories characterizing ASD

abilities [9], [14]: since the maturation of perceptual abilities is delayed in ASD, we could assume that cross-modal connections are more negative in ASD children of 6-7 years of age compared to their NT peers.

As it is evident from Fig. 3, the model provides a good fit with empirical data of [2].

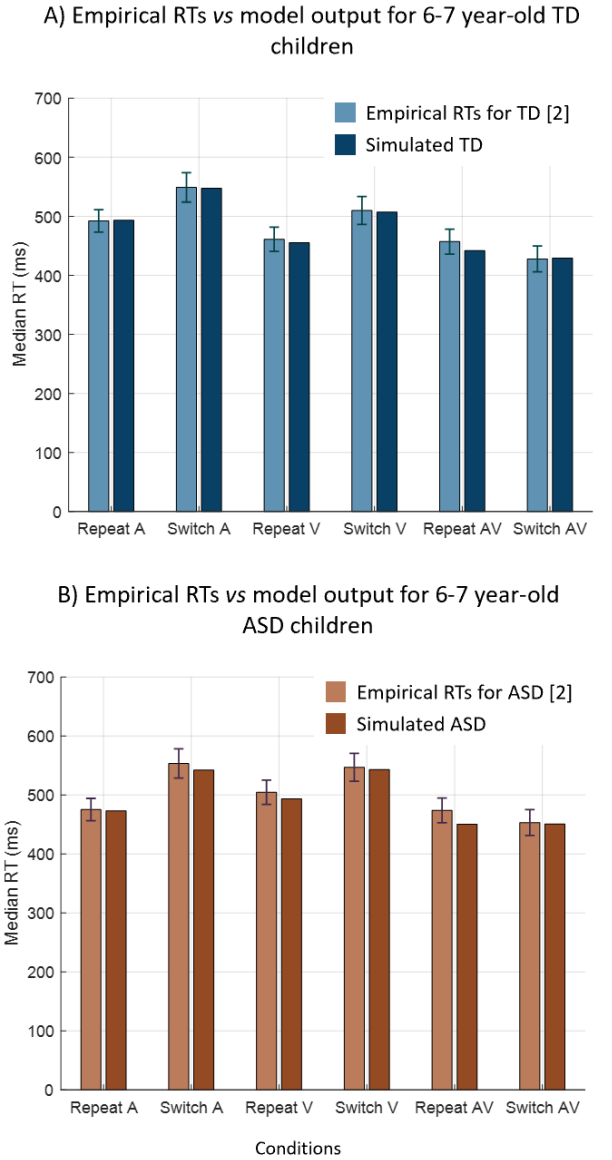


Fig. 3: Simulated RTs vs Median RTs of the subjects' population. For each stimulus configuration (A and V Repeat, A and V Switch, AV Repeat and Switch) RTs obtained with the model were compared with the experimental median RTs extracted from the population of real subjects. To set the basal configuration of parameters, no noise was added (neither on the input, nor on the synapses).

### B. Model's behaviour

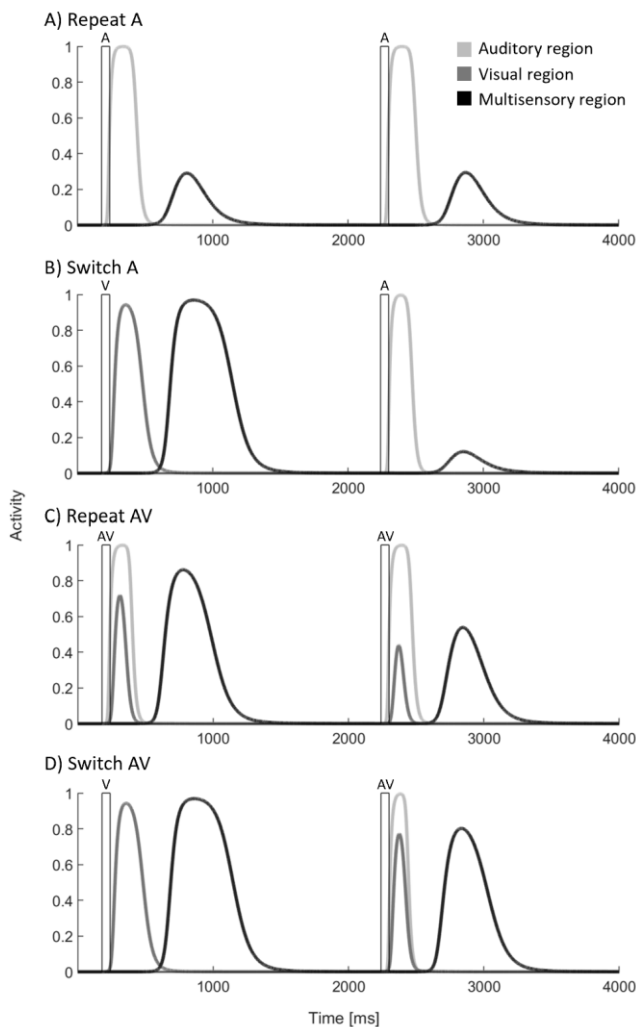
To clarify how each implemented mechanism contributes to the model's behaviour, in Fig. 4 we report the activities elicited in the NT network's regions in four exemplary conditions of unisensory (A) and multisensory (AV) repeat and switch.

Comparing the auditory repeat and switch conditions (Fig. 4A and 4B), we can notice that the activation elicited in the model by the second stimulus (A) is lower in the switch case. This is due to the inhibitory effect exerted by the prior input (V): the visual inhibitory interneuron, excited by the preceding visual stimulus, inhibits the auditory input region through the

feedback projections, characterized by a slow dynamic (i.e. the effect acts over a long period of time). The consequent lower activity produces a delayed activation of the multisensory region, resulting in longer RT.

Similar results were found for visual repeat/switch trials.

In case of multisensory stimulation, in addition to this slow inhibition that the prior input exerts on the processing of the current AV stimulus, the A and V components of the AV input reciprocally compete, by means of the direct inhibitory cross-modal synapses. This results in a lower activity elicited by an AV stimulus in the input layer (Fig. 4C and D) compared to the case of unisensory stimulation. This inhibition counterbalances the integration occurring in the M region, mediated by the excitatory feedforward projections. The presence of cross-modal inhibitory synapses is critical to explain the behavioural data. Indeed, without this mechanism a cross-modal input would produce RTs much faster than those found in [2].



**Fig. 4:** Example of the network behavior, in the case of A) Repeat A, B) Switch A, C) Repeat AV and D) Switch AV conditions, for NT children. Rectangular wave represents the external visual stimuli presented with an ISI of 2000ms. Light grey lines describe the activation of the auditory region, dark grey lines describe the activation of the visual area. Black lines represent the activity elicited in the output region in response to this stimulation.

#### IV. CONCLUSION

In this work, we developed a biologically plausible neurocomputational model of perceptual abilities and sensory

interactions in children. This model should be thought of as a “engineering” tool, to test and interpret current empirical knowledge, elucidating the neural mechanisms in play.

In the first part of our work, we investigated the nature of cross-modal interactions in children. Empirical evidence [1] identified a default state of cross-modal competition in kittens, which could potentially represent a mechanism preserved also in human brain. Indeed, recent children’s behavioural data [2] are consistent with such a competitive architecture. In our model, we linked these empirical results with a plausible neural implementation, in which the default competition is realized via inhibitory cross-modal projections.

In the second part, we tested how altered perceptual mechanisms could contribute to ASD phenotype and we suggested that the slower maturation of perceptual abilities in ASD would result in a stronger cross-modal competition in these children compared to their NT peers.

Overall, the importance of this work is that it provides with a deeper insight into perceptual function, dysfunction and behaviour in children. One field expected to benefit from this in the near future is the neurorehabilitation: by elucidating the underlying neural mechanisms, plausible neural networks will allow to make predictions, design specific interventions and explain their effects.

#### REFERENCES

- [1] L. Yu, C. Cuppini, J. Xu, B. A. Rowland, and B. E. Stein, “Cross-modal competition: the default computation for multisensory processing”, *J Neurosci.*, vol. 39(8), pp. 1374–1385, 2019.
- [2] M. J. Crosse, J. J. Foxe, K. Tarrit, E. G. Freedman, and S. Molholm, “Resolution of impaired multisensory processing in autism and the cost of switching sensory modality”, *Commun Biol.*, vol. 5(1), pp. 601, 2022.
- [3] M. A. Meredith and B. E. Stein, “Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration”, *J Neurophysiol.*, vol. 56(3), pp. 640–662, 1986.
- [4] B.E. Stein and M.A. Meredith, *The merging of the senses*. Cambridge, MA: Massachusetts Institute of Technology, 1993.
- [5] M. A. Meredith, J. W. Nemitz, and B. E. Stein, “Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors”, *J Neurosci.*, vol. 7(10), pp. 3215–3229, 1987.
- [6] M. T. Wallace, T. J. Jr Perrault, W. D. Hairston, and B. E. Stein, “Visual experience is necessary for the development of multisensory integration”, *J Neurosci.*, vol. 24(43), pp. 9580–9584, 2004.
- [7] C. Cuppini, M. Ursino, E. Magosso, M. J. Crosse, J. J. Foxe, and S. Molholm, “Cross-sensory inhibition or unisensory facilitation: A potential neural architecture of modality switch effects”, *J. Math. Psychol.*, vol. 99, pp. 102438, 2020.
- [8] B. de Gelder, L. Vroomen, and L. Van der Heide, “Face recognition and lipreading in autism”, *Eur. J. Cogn. Psychol.*, vol. 3, pp. 69–86, 1991.
- [9] J. J. Foxe et al., “Severe multisensory speech integration deficits in high functioning school-aged children with autism spectrum disorder (ASD) and their resolution during early adolescence”, *Cereb Cortex.*, vol. 25(2), pp. 298–312, 2015.
- [10] R. A. Stevenson et al., “Multisensory temporal integration in autism spectrum disorders”, *J. Neurosci.*, vol. 34, pp. 691–697, 2014.
- [11] R. A. Stevenson et al., “Evidence for diminished multisensory integration in autism spectrum disorders”, *J Autism Dev Disord.*, vol. 44, pp. 3161–3167, 2014.
- [12] R. A. Stevenson et al., “The cascading influence of multisensory processing on speech perception in autism”, *Autism*, vol. 22, pp. 609–624, 2018.
- [13] A. B. Brandwein et al., “The development of multisensory integration in high functioning autism: high-density electrical mapping and psychophysical measures reveal impairments in the processing of audiovisual inputs”, *Cereb Cortex*, vol. 23, pp. 1329–1341, 2013.
- [14] N. Taylor, C. Isaac, and E. Milne, “A comparison of the development of audiovisual integration in children with autism spectrum disorders and typically developing children”, *J Autism Dev Disord.*, vol. 40, pp. 1403–1411, 2010.