Minimalistic Toy Robot Encourages Verbal and Emotional Expressions in Autism

Irini Giannopulu  
Pierre & Marie Curie University  
Cognitive Neuroscience  
Paris, France  
email:igiannopulu@psycho-prat.fr

Valérie Montreynaud  
Center of Medical Psychiatry  
Paris, France  
email:v.montreynaud@gpspv.fr

Tomio Watanabe  
Okayama Prefectural University  
Department of Systems Engineering  
Okayama, Japan  
email:watanabe@cse.oka-pu.ac.jp

Abstract—Language offers the possibility to transfer information between speaker and listener who both possess the ability to use it. Using a “speaker-listener” situation, we have compared the verbal and the emotional expressions of neurotypical and autistic children aged 6 to 7 years. The speaker was always a child (neurotypical or autistic); the listener was a human InterActor or an InterActor robot, i.e., a small toy robot that reacts to speech expression by nodding only. The results suggest that a robot characterized by predictable reactions facilitate autistic children in expression. When comparing to the performance of neurotypical children, the data would indicate that minimalistic artificial environments have the potential to open the way for neuronal organization and reorganization with the ability to support the embrainment of verbal and emotional information processing among autistic children.

Keywords—brain development; neurotypical children; children with autism; minimalistic robot; language; emotion; listener-speaker.

I. INTRODUCTION

Development is the result of a complex process with at least three foci, one in the central nervous system, one in the mind and one in the child’s dynamic interactions with the natural vs. artificial environment, that is, robots. The human child brain undoubtedly has its own dynamics (probably because of the extensive expression of genes in the brain) that allows neurons to interact, create their own multimodal nature, which in turn, affects the nature, development and function of the brain areas [1]. Verbal and nonverbal cognition, as well as emotion develop at the interface between neural processes. Toys have a central role. Children tend to play with toys in the first year of life. Toys are put together, sucked, kissed, piled on top of each other. The ability to play with toys becomes more and more sophisticated as development progress. The young children start to play with the toys in a symbolic manner by pretending that the toys represent something else that they love or not. They are able to move on from using themselves as an agent to using toys as (active) agents and carry out various actions [2]. Toys seem to provide an interesting account of “how” physical objects are able to act as support for the symbolic play of children. Symbolic play, like verbal development, emerges progressively as toys are the indices that assist the child to go in [3]. With development, symbolic play with action grows into language. With that in mind, we can imagine a scenery of communication between two persons: one speaking the other listening.

Neurotypically developing listeners and speakers are able to consider verbal, nonverbal (i.e., head nods) emotional (i.e., facial expressions) conventions and rules as well as each other when making referential statements. This is potentially due to the formation of a neural multimodal network, which naturally follows the evolution of the brain [4]. Using a modeling approach, recent neuroimaging studies have reported that both speech comprehension and speech expression activate a bilateral fronto-temporo-parietal network in the brain, fully characterized by the dynamic interaction among all the components [5]. Different studies emphasize the importance of multiple cortical (e.g., prefrontal cortex, temporal and parietal cortices) and subcortical areas (e.g., basal ganglia, hippocampus and cerebellum) not only for production and reception of speech but also, for cognitive nonverbal and emotional processes [1][6].

Failure of the exterior superior temporal sulcus [7], of the interior temporal lobe, amygdala included [8], of the connectivity between temporal regions [9], as well as of the inferior prefrontal cortex [10], i.e., the mirror neurone system, is accepted as an explanation for atypical neurodevelopment, such as autism [11][12]. The atypical neural architecture causes impairment in social interaction, in communication skills and interests [13][14][16] and reduces the ability of mentalizing, i.e., making representations based on the referential statements of other people [17].

Autistic children listeners and speakers perform less well than neurotypical children in conversation especially when the listener is a human, (human is essentially characterized by a high degree of variability on verbal and nonverbal emotional reactions, i.e., unpredictable reactions [18]). Adding the fact that the child is impaired in interpreting the referential statements of other people [12], the listener’s verbal and nonverbal emotional contributions are not always scrutinized. There are at least two main reasons for this. The first reason is associated with the fact that autistic children have continual comprehension and language expression problems. Even if autistic children acquire language, it is often lacking any depth and is characterized by a paucity of imagination.
[19]. The second reason is that autistic children experience difficulties in perception and emotion, functions which are linked to language [20] and also to social interaction and mentalising [9].

Trying to analyze child-robot interaction, different approaches have been developed. Different approaches have shown that animate robots using different stimulation encourage interaction in autistic children [20]. Quantitative metrics for autism diagnosis and treatment including robots have been developed [21]. Despite these studies, only marginal attention has been paid to the comparison of neurotypical and autistic children in human-human and human-robot interaction. Using a “speaker-listener” situation, we have compared the verbal and emotional expressions of neurotypical and autistic children aged 6 to 7 years. The speaker was always a child (neurotypical or autistic); the listener was a human or an InterActor robot, i.e., a robot, which reacts to speech expression by nodding only. Given the fact that the InterActor robot is characterized by a low degree of variability in reactions (i.e., predictable reactions) and the human by a high degree of variability in reactions (i.e., unpredictable reactions), our general hypothesis is that verbal and emotional expressions of autistic children could better be facilitated by the InterActor than by the human.

Beginning with the design of the study, we will continue with the analysis of the results in both neurotypical and autistic children. Then, we will discuss the importance of minimalistic artificial environments as support for the embodiment of cognitive verbal and emotional information processing.

II. METHOD

A. Participants

Two groups of children, one “neurotypical” and one “autistic” participated in the study. Twenty neurotypical children (10 boys and 10 girls) composed the “neurotypical group”; twenty children (14 boys and 6 girls) composed the “autistic group”. The developmental age of typical children ranged from 6 to 7 years old (mean 6.1 years; standard deviation 7 months). The developmental age of autistic children ranged from 6 to 7 years old (mean 6 years; standard deviation 8 months). Their mean age when first words appeared was 28 months (standard deviation 7 months). The autistic children were diagnosed according to the DSM IV-TR criteria of autism [22]. The Childhood Autism Rating Scale CARS [23] has been administrated by an experienced psychiatrist. The scores varied from 31 to 35 points signifying that the autistic population was composed of mild-moderate children with autism. They were all verbal. All autistic children were attending typical school classes with typical educational arrangements. The study was approved by the local ethics committee and was in accordance with the Helsinki convention. Anonymity was guaranteed.

B. Material

• Robot

An InterActor robot, i.e., a small toy robot, called “Pekoppa”, was used as a listener [24]. Pekoppa is shaped like a bilobed plant and its leaves and stem make a nodding response based on speech input and support the sharing of mutual embodiment in communication (Figure 1). It uses a material called BioMetal made of a shape-memory alloy as its driving force. The timing of nodding is predicted using a hierarchy model consisting of two stages: macro and micro (Figure 2). The macro stage estimates whether a nodding response exists or not in a duration unit, which consists of a talkspurt episode $T(i)$ and the following silence episode $S(i)$ with a hangover value of 4/30 s. The estimator $Mu(i)$ is a moving-average (MA) model, expressed as the weighted sum of unit speech activity $R(i)$ in (1) and (2). When $Mu(i)$ exceeds a threshold value, nodding $M(i)$ also becomes an MA model, estimated as the weighted sum of the binary speech signal $V(i)$ in (3). Pekoppa demonstrates three degrees of movements: big and small nods and a slight twitch of the leaves by controlling the threshold values of the nodding prediction. The threshold of the leaf movement is set lower than that of the nodding prediction.
Prediction. The threshold of the signal whether a nodding response exists or not in a communication (Figure 1). It uses an InterActor robot, called “Pekoppa”, which also becomes an MA model, estimated as (1) and (2). When (1) exceeds a threshold value, nodding is predicted using a hierarchy model (Figure 2). The macro stage estimates the regularity of one talkspurt duration in the i-th duration unit $T(i)$ in (3). Pekoppa demonstrates three $i$ and the following silence episode $S(i)$ in (3).

$$M_u(i) = \sum_{j=1}^{i} a(j)R(i-j) + u(i)$$  \hspace{1cm} (1)

$$R(i) = \frac{T(i)}{T(i) + S(i)}$$  \hspace{1cm} (2)

$$M(i) = \sum_{k=1}^{K} b(j)V(i-j) + w(i)$$  \hspace{1cm} (3)

$\ a(j)$: linear prediction coefficient
$\ T(i)$: talkspurt duration in the i-th duration unit
$\ S(i)$: silence duration in the i-th duration unit
$\ u(i)$: noise

The inter-condition interval was approximately about 30 minutes. For each child, the whole experimental session lasted 15 minutes (Figure 3).

Procedure

For both groups, the study took place in a room, which was familiar to the children. We defined three conditions: the first one was called “rest condition”; the second was named “with human” (child-adult) and the third one was called “with robot” (child-Robot, i.e., child-Pekoppa). The second and third conditions were counterbalanced across the children. The duration of the “rest condition” was 1 minute; the second and third conditions each lasted approximately 7 minutes.

The inter-condition interval was approximately about 30 seconds. For each child, the whole experimental session lasted 15 minutes (Figure 3).

Figure 3. Listener-Speaker Situation

In order to neutralize a possible “human impact” on children’s behavior, the experimenter was the same person for each child in each condition and group (“neurotypical” and “autistic”). At the beginning of each session, the experimenter presented the robot to the child explaining that the robot nods whenever the child speaks. Then, the experimenter hid the robot. The session was run as follows: during the “rest condition”, the heart rate of each child was measured in silence. At the end of that condition, the child was also asked to estimate the intensity of her/his own emotion on a scale ranging from 1 (the lowest intensity) to 5 (the highest intensity) [25][26]. During the “with human” condition, the child was invited to discuss with the experimenter. The experimenter initiated discussion and after listened to the child acting as the speaker. The heart rate, as well as the frequency of words and verbs expressed by each child was measured. During the “with robot” condition, Pekoppa was set to nod movements; the experimenter gave the robot to the child inviting the child to use it. The robot was the listener, the child was the speaker and the experimenter remained silent and discreet. The heart rate and the frequency of words and verbs expressed by the child was recorded once again. At the end of the session, the child was invited to estimate the intensity of its own emotion on the same aforementioned scale. At the end of the experiment, each child was invited to respond to two questions: 1) how did you find Pekoppa? 2) did you enjoy yourself with Pekoppa? [see also 26].

Analysis

The analysis was based on the following dependent variables a) the heart rate measured in beat per minute (bpm) b) the number of nouns and verbs expressed by each child and c) the intensity of emotional feeling (auto-estimation of emotion). The data analysis was performed with SPSS Statistics 17.0 [27].

III. Results

The distributions of heart rate, words and emotional feeling reported in both age groups approximate a parametric shape. With such distributions, the mean has been chosen as a central index for comparisons. We performed statistic of comparisons using the t-student test, the ANOVA’s test and the chi-square test to examine differences in heart rate, number of words and intensity of emotional feeling between the two experimental conditions (“with human” and “with Robot” i.e., Pekoppa), for neurotypical and autistic children. The obtained results were very similar. We present the results of chi-square test ($\chi^2$ test), which can be used as a substitute for t and ANOVA tests [28].

Figure 4 represents the mean heart rate of neurotypical and autistic children both at the inter-individual and the intra-individual levels.

At the intra-individual level, the statistical analysis showed that relative to the “rest condition”, the mean heart rate of neurotypical children was higher when the children were in contact with the InterActor robot ($\chi^2$=6.68, $p$=0.01) than when they were in contact with the human ($\chi^2$=4.09, $p$=0.05). However, the mean heart rate of neurotypical children didn’t differ significantly when they interacted...
with the human or with the InterActor robot ($\chi^2=2.83$, $p>0.05$). Similarly, relative to the “rest condition”, the mean heart rate of autistic children was higher when they interacted with the InterActor robot ($\chi^2=7.01$, $p<0.01$) than when they interacted with the human ($\chi^2=5.01$, $p<0.05$). Finally, the mean heart rate of autistic children was higher when they were with the InterActor robot than when they were with the human ($\chi^2=7.84$, $p<0.01$).

Two independent judges unfamiliar with the aim of the study completed the analysis of the number of words for each child in each experimental condition (“human InterActor” and “robot InterActor”). Both performed the analyses of audio sequences. Inter-judge reliability was assessed using intra-class coefficients to make the comparison between them. The inter-judge reliability was good (Cohen’s kappa=0.82).

At the inter-individual level, as shown in Figure 5, the mean number of words (nouns and verbs) was low in the “with human” condition for autistic children ($\chi^2=4.86$, $p<0.05$) and in the “with robot” condition for neurotypical children ($\chi^2=5.98$, $p=0.025$). The mean number of words expressed by autistic children in the “with robot” condition didn’t differ from the mean number of words expressed by neurotypical children in the “with human” condition ($\chi^2=1.34$, $p>0.10$). At the intra-individual level, the mean number of words was higher when the autistic children had the robot as interlocutor than when the interlocutor was a human ($\chi^2=5.97$, $p<0.025$). The quasi opposite configuration was observed for the neurotypical children ($\chi^2=4.78$, $p<0.05$).

![Figure 4. Mean Heart Rate](image1)

![Figure 5. Number of words (nouns & verbs)](image2)

IV. DISCUSSION

The present study aims at analyzing the embrainment of verbal and emotional expressions in neurotypical and autistic children aged 6 to 7 years. Our approach centered on investigating the effects of a human or an InterActor robot in the context of a “speaker-listener” situation: the speaker was always the child; the listener was a human or an InterActor robot, i.e., Pekoppa. To this end, physiological data (i.e., heart rate), as well as behavioral data (i.e., number of nouns and verbs in addition to the emotional feeling reported) were considered. The results showed that 1) the heart rate of autistic children is low when the listener was a human and increased nearer to levels of neurotypical children when the listener was the InterActor robot; 2) the number of words expressed by the autistic children was higher when the interlocutor was the
robot; 3) the emotional feeling reported increased after the interaction with the InterActor robot.

Fundamentally, the results are consistent with our hypothesis according to which the predictability of the InterActor robot would facilitate the emotional and verbal expressions of autistic children. Our results showed significant differences of heart rate depending on whether the listener was a human or a robot. When the listener was a human, the children showed a low heart rate; when the listener was an InterActor robot, their heart rate increased. Such a result cannot be attributed to an order effect as the order of “human-human” and “human-robot” conditions have been counterbalanced. On the contrary, it could be understood as an effect of the InterActor robot on autistic children’s mental state. This interpretation is supported by the fact that when the autistic children had the InterActor robot as listener, their heart rate didn’t differ from the heart rate of neurotypical children in the same condition. It is also interesting to note that the heart rate of the neurotypical children didn’t differ when the listener was a human or a InterActor robot. The observation reported above could reveal that an InterActor robot might improve autistic children behavior. This inference is reinforced by the fact that the physiological data we recorded reflects the modifications of orthosympathetic and parasympathetic autonomous nervous system, which is dynamically (and bidirectionally) connected to the central nervous system [29][30]. Physiologically, the lower regulation of heart rate (in “with human” condition) reflects poorer action of the myelinated vagus nerve [31], which in turn would signify poor neural activity in temporal cortex (amygdala included), in cingulate cortex and in prefrontal cortex [32][33]. This neural architecture is hypo-activated in children with autism [14][15], causing impairment in cognitive verbal, nonverbal and emotional behavior [16][20]. Such hypo-activation would explain autistic children’s behavior when the listener is the human. A restricted number of studies exist on the evaluation of heart rate in autistic children. Some of them have reported that autistic children display lower heart rate than typically developing children [34][35], some others have found the opposite [36][38]. The aforementioned studies suggest that autistic children show disruptions in autonomic responses to environmental (human) stressors. Methodological problems associated to the limited number of autistic children, to their clinical heterogeneity as well as to the various procedures and measures used are on the basis of the opposing reported data. None of the above studies have evaluated autonomic activity in a robot-child interaction vs. human-human interaction. To our knowledge, the present study is the first one having a homogeneous group of mild-moderate children with autism, which is matched on developmental age with a group of typically developing children. Contrary to the previous studies, our findings indicate that not only are there no disruptions in autonomic responses but that these responses don’t exceed the physiological limits. Apparently, when the listener is the InterActor robot, the heart rate of children with autism increases likely indicating a “mobilisation” of a given mental state. Such “mobilisation” would provide support for the social engagement of autistic children. Namely, by making the autistic children available to engage emotionally (and verbally), the InterActor robot seems to modify their neural activity: the children would enjoy participating. It is noteworthy that they also verbalized such pleasurable sentiments at the end of the experiment. Such information is presented here below. Essentially, the present results are consistent with our previous assumptions following which toy robots would improve autistic children brain functioning [26][39][40].

The above considerations could account for the number of words (nouns and verbs) expressed by the children. Even if the autistic children were verbal, the present finding indicated that when the listener was an InterActor robot, the number of words expressed by the autistic children was higher than the number of words they express when the listener was a human. Interestingly, such verbal behavior doesn’t differ from that of neurotypical children when these latter had a human as listener. Once again, the use of the InterActor robot seems to afford autistic children the ability to express themselves as neurotypical children do with humans. These data are consistent with previous studies, which have demonstrated that verbal expression can be facilitated by the active (but discreet) presence of a robot [20].

Although neurotypical children didn’t report emotional feeling changes after their interaction with the robot, autistic children said to feel better after interaction with the robot. This is coherent not only with the physiological data we observed (heart rate) but also with parent accounts. At the end of the experiment, many parents announced: “s/he is happy”, “s/he likes your robot”. Autistic children also conceded that the robot was “cute”, “cool” “genius”, some of them even said: “if I had the robot, I would talk to it all the time”. Some of them imitated the robot verbally (and emotionally).

It could be argued that the “autistic group” was made up of verbal children and that the results we observed might be due to the actual verbal capabilities of the children. However, we must underline that these children expressed themselves (both emotionally and verbally) only when the listener was the InterActor Robot. Although our results are statistically significant, we recognize that the size of our group is limited to twenty children only. We aim to study the behavior of other age groups as well. Finally, it is obvious that what we need to develop is a follow-up study to prove that the InterActor robot is the robot, which can sustainably improve the emotional and verbal behavior of autistic children.

V. CONCLUSION AND FUTURE WORK

Given the present findings, it can be concluded that an InterActor robot characterized by small-variance nonverbal behavior, (i.e., nodding when children speak), facilitates verbal and emotional expressions of autistic children. This might be related to autistic children preferences. Autistic
children are rather interested in minimalist objects to which they can assign mental states of their own or of others [41]. Such a behavior might be interpreted as reflecting the children’s willingness to communicate with humans using the robot: the InterActor toy robot is a miniature of a human listener, i.e., the autistic children can handle the head nods (as neurotypical children do with humans). These results (consistent with previous studies) appear to indicate that minimalist artificial environments could probably be considered as the root of neuronal organization and reorganization with the potential to improve brain activity in order to support the embrainment of cognitive verbal and emotional information processing [41][44].

Additional studies are required with typical and autistic children. Longitudinal follow-up of the same children is necessary to examine the efficiency of minimalist robots in improving the activity of autistic children. This is what we're developing currently in France and in Japan. In addition, with a new study we analyse the enroachment [45] in conscious and unconscious level in neurotypical children aged 6 and 9 years old.

ACKNOWLEDGMENT

To all the participants and their parents, the Major, the Pedagogical Inspector, the Educational Advisor, the Director and the team of principal elementary school of the first district of Paris, and the National Department of Education and Research. The research is supported by the Franco-Japanese Foundation.

REFERENCES

27. SPSS STATISTICS 17.0, Python Software Foundation, 2008.