STUDYING AUTISM THROUGH THE NAO ROBOT
THE ARTS RESEARCH COLLABORATORY
FACULTY OF ARTS, RYERSON UNIVERSITY

CONVERTING EXISTING VOCABULARY TESTS INTO ROBOTIC PROGRAMS

Does robotic interaction provide a new context, enhancing response accuracy in children with ASD performing cognitive representational tasks related to lexical semantic categories?

Stéphanie Walsh Matthews, Alexandra Marquis, Saijal Suri, Jamin Pelkey

Ryerson University

Why HRI to Study Semantic Organization?

Human **R**obot **I**nteraction (HRI) is a proven and effective way to engage children with **A**utism **S**pectrum **D**isorder (ASD). Humanoid robotics provide a reduced yet dynamic interface that is predictable and engaging. With it, a number of socio-communicative activities can be carried out and observed.

Recent studies (e.g., Walsh Matthews et al., 2016; Kim et al., 2015; Shamsuddin et al., 2012) show increased social communicative behaviour by children with ASD following HRI. To further these test effects, we designed a study using the NAO robot that would substitute human-conducted vocabulary testing with HRI by re-creating Tager-Flusberg's (1985) experiment.

Proposed Study: from PPVT to RVT

In Tager-Flusberg's (1985) article, entitled "The Conceptual Basis for Referential Word Meaning in Children with Autism" two studies propose examining the representational nature of word meaning in children with ASD and to compare results obtained to matched typical and atypical children. In the first experiment, TF uses the Peabody Picture Vocabulary Test (PPVT) to assess the understanding of words "for basic and superordinate level categories" (1169) to better understand the organization of semantic representation in children with ASD.

Working Hypotheses

Our study aims at simulating the same experiment using our designed robotic behaviours on a NAO robot to mediate the test. By programming a Robotic Vocabulary Test (RVT) to examine the nature of substantive lexical semantic representations in children with ASD, we:

1) Verified if **humanoid robotics** could be used **to conduct standard vocabulary tests** by comparing results to that of Tager-Flusberg (1985);

2) Investigated if **HRI would increase the accuracy of responses** based on previous studies reporting that HRI did in fact increase socio-communicative expressions in children with ASD;

3) Pursued our previous working hypothesis in **Cognitive Linguistics** investigating **Idealized Cognitive Models** and role reversal when the **robot malfunctioned** *in situ* (non-controlled).

Methods

Consent and Schedule	Ryerson University's Research Ethics Board approved the current study and consent was collected from caregivers and carried out during the child's regular schedule for IBI or ABA.
RVT	Child's Therapist was present during the test. Robotic-Vocabulary Test was conducted by a Card Facilitator (directly involved), a Robot Technician, and two Observers (not directly involved).
Mixed Approach	Audio recorded sessions, ethnographic observation notes, data tabulated notes for RVT Scores by at least two Observers.
Child Language Data Exchange System (CHILDES)	Sessions were transcribed by at least two independent researchers using CHILDES techniques.
Measures of Performance	*Percent Affirmative scores =* # of times the child answered yes/number of cards *Percent Correct scores =* # of accurate responses/number of cards
MLU of non-RVT utterances	Mean Length Utterances (MLU) was calculated by counting the number of morphemes and dividing it by the number of utterances spoken by the child during non-RVT prompts (MacWhinney, 2000)

From PPVT to RVT

By replacing the PPVT with a programmed RVT version, ROBOT asks CHILD if card is a match for its related category. CHILD states "yes" or no" and answer is recorded onto a data sheet.

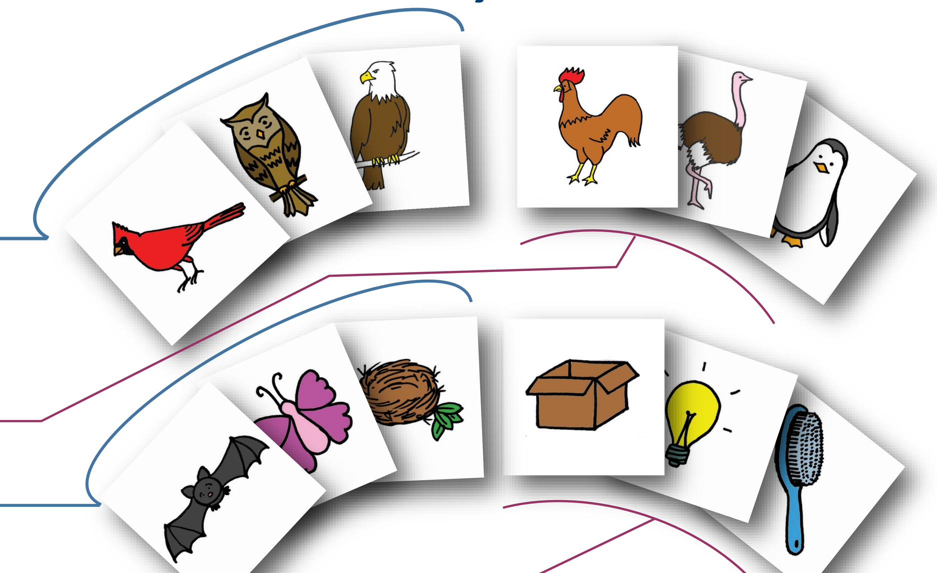
Basic Level Categories

This category includes BIRDS and BOATS.

Objects in this category **have observable features that resemble** in perceptual type.

They are organised by:

- **Central Members**
- **Peripheral Members**
- **Related Foils**
- **Unrelated Foils**



Superordinate Level Categories

This category includes more **abstract** examples of **concept categories** including FOOD and TOOLS.

These are **not grouped** according to perceptual characteristics. They are organised by:

- **Central Members**
- **Peripheral Members**
- **Related Foils**
- **Unrelated Foils**

RVT Programming

```
        pass

def onInput_onStart(self):
    #reset retry count
    global script
    ALMemory.insertData("peabodycardsgame/retrycount", 1)
    if(self.questionNO <= 10):
        if (self.questionNO % 3 == 0):
            Question = script['72']
        else:
            Question = script['18']

    else:
        Question = script['17']

    self.questionNO+=1
```

Peabody Question

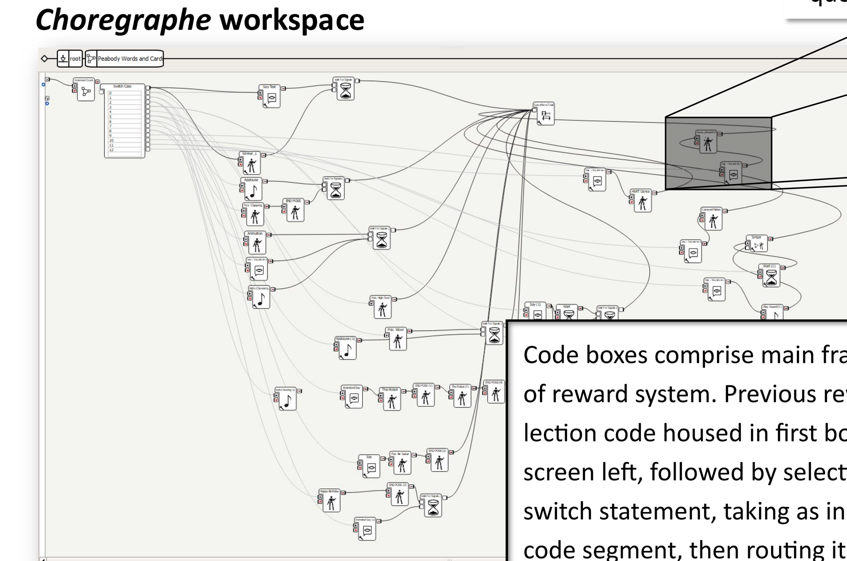
Code segment selecting correct question from NAO to user. Data read in from pre-existing "script" file with list of images on associated image cards in prescribed order. Appropriate verbalization selected from script array at appropriate index allows structured repetitive questions creating highly predictable environment.

```
def onLoad(self):
    #~ puts code for box initialization here

    pass

def onUnload(selfs):
    #~ puts code for box cleanup here
    pass

def onInput_onStart(self):
    if (self.nInterstitial % 10 == 0):
        rewardNum = self.interCounter
        self.interCounter+=1
        if(self.interCounter==13):
            self.interCounter = 9
        self.log("interstitial reward")
        self.log(rewardNum)

    else:
        rewardNum = random.randint(0,8)
        self.log("normal reward")
        self.log(rewardNum)
    self.nInterstitial+=1

    #reset retry count
    try:
        self.Memory.insertData("peabodycardsgame/retry
```

Interstitial Reward

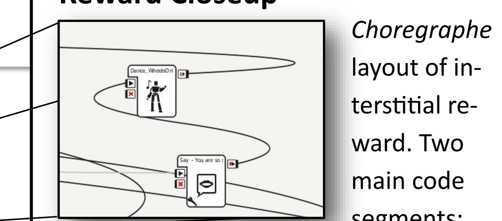
Code segment selecting appropriate reward to follow successful completion of question by user. Two reward types coded: (1) short song/ dance or robot interaction; (2) longer song/ dance segment (up to 40 sec.). Interstitial rewards run every 10 questions to elicit user interaction and provide break from PPVT. Songs selected for likely familiarity. Rewards selected every tenth question.

***Choregraphe* workspace**



Code boxes comprise main framework of reward system. Previous reward selection code housed in first box at screen left, followed by selection box switch statement, taking as input the reward number selected by the previous code segment, then routing it to appropriate reward boxes in timeline web.

Reward Closeup

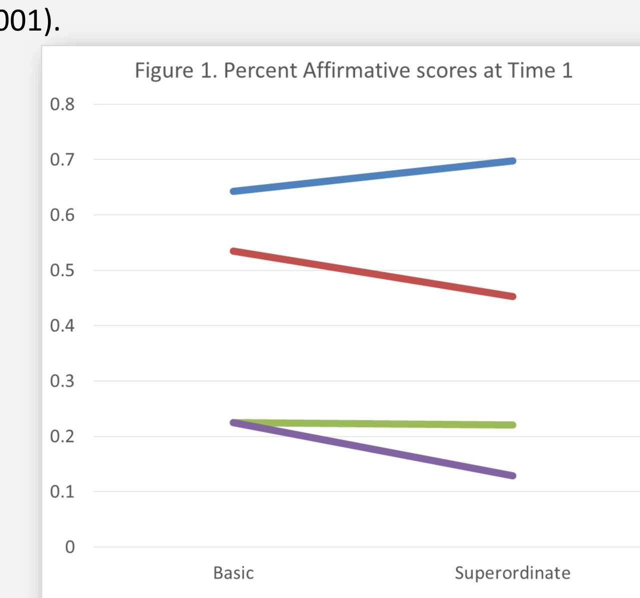


Choregraphe layout of interstitial reward. Two main code segments: (1) dialogue box instructing NAO to provide encouragement to user; (2) code necessary for robot to activate reward-specific music and dance routine. Dialogue box: "Yay! You're a pro! Sing along with me!" Song/Dance routine: "Wheels on the Bus".

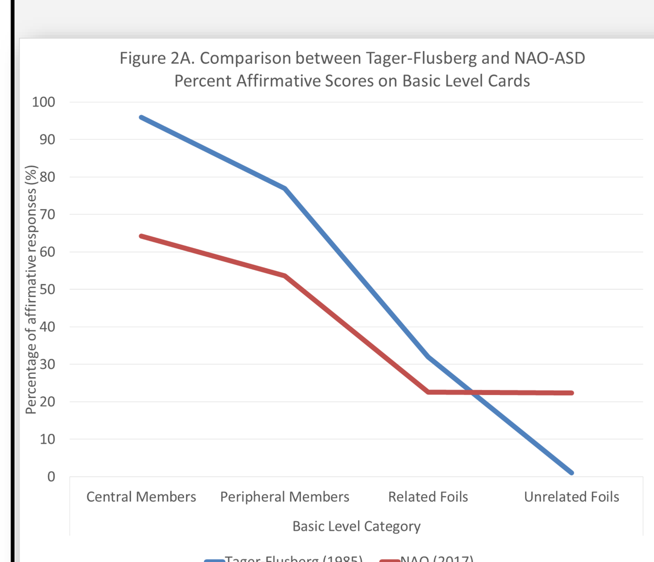
Results

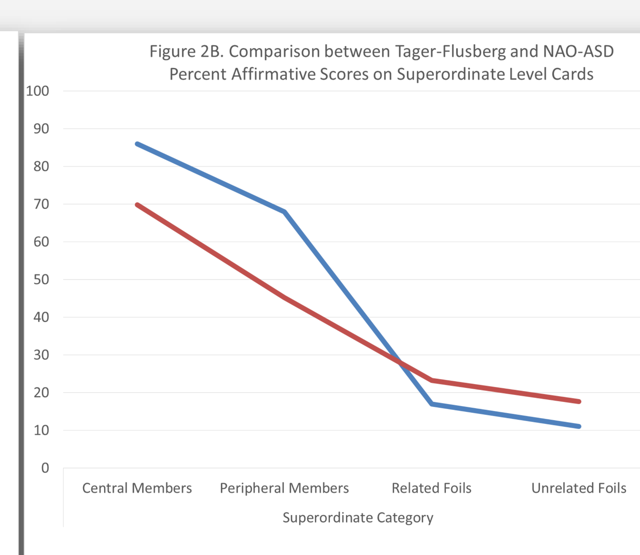
1. Humanoid robot-led vocabulary test compared to Tager-Flusberg's (n = 12)

Similar to Tager-Flusberg's (1985) Peabody Vocabulary Test, we found a significant main effect of Stimuli Type (4: Central, Peripheral, Related, Unrelated) on Percent Affirmative Scores at Time 1 (F(3, 39) = 25.36, p < .001).



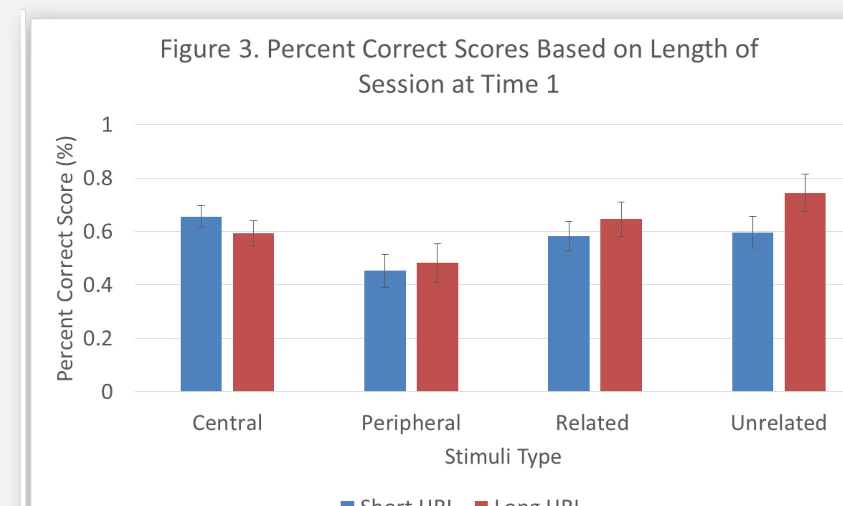
Comparison between Tager-Flusberg's (1985) Peabody Vocabulary Test Findings and the NAO-ASD Findings on Percent Affirmative Scores. The scores did not significantly differ, t (14) = .57, p = .58





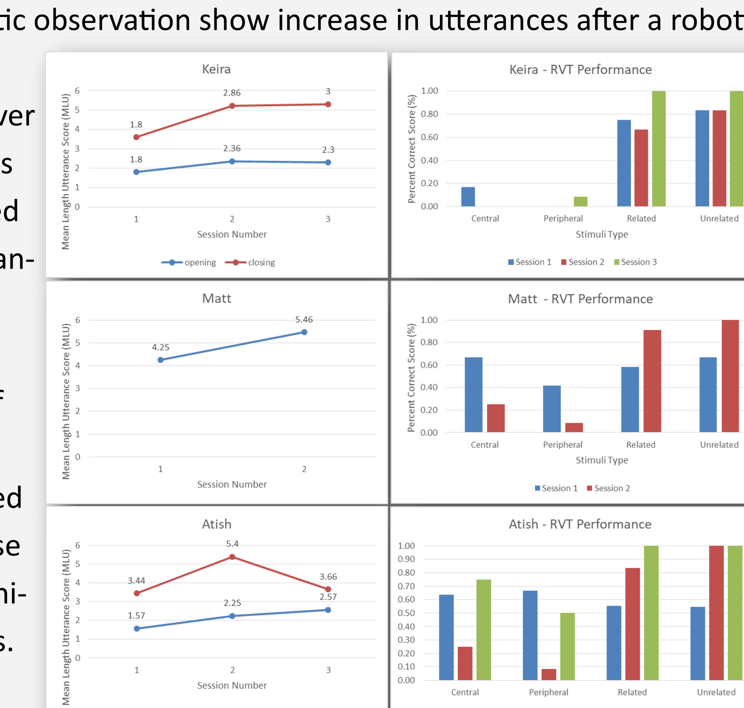
2. Does HRI increase the accuracy of responses?

Comparison between children who had short HRI sessions with those who had long HRI sessions at Time 1 on their performance. Repeated Measures ANOVA indicated no significant main effect of Stimuli Type or Length of Session (n = 12)



3. ICM Results

- MLU findings continue to support increase in utterances production for participants during sessions and over-time.
- MLU findings and contextual linguistic observation show increase in utterances after a robotic malfunction.
- Changes in anthropomorphic bias over time (exposure to HRI) also increases peer responsiveness and is expressed through ICM of peer and therapist language models.
- Although repeated measure ANOVA indicate no significant main effect of stimuli type on length of session, increased MLU over time and increased chance of malfunctions does increase language productivity based on cognitive modeling during longer sessions.



Findings

HYPOTHESIS 1: Can you do PPVT as RVT?

Yes. The RVT is more repeatable and highly randomized. It includes a child specific rewards system and provides low impact testing.

HYPOTHESIS 2: Would a RVT yield improved results over the PPVT thanks to HRI?

RVT provides **similar results** for central and basic superordinate categories as TF's 1985 PPVT. During the observed RVT session, children did answer "yes" more often for incorrect answers and false positives, thus reducing the test accuracy. RVT results over time did show **slight improvement** for certain medium- proficiency level speakers.

HYPOTHESIS 3: Do robot malfunctions invite novel speech acts or increased utterances?

Yes. A number of case studies from low, medium, and high proficiency level speakers show an increase in utterance production and MLU when the robot did malfunction. In addition, for participants with medium to high-proficiency language levels, robotic-error produced **role reversals**. First reversal trends towards the ICM of /therapist/ and, second and subsequent reversals, included the ICM of /peer/. The peer model **increased sociability** and **MLU**, as well as sparked a series of questions engendering the **anthropomorphic bias.**

References

Danesi, M. (2006). The Sense implication hypothesis and idealized cognitive models Semiotica. 161 (1/4): 185-198.
Chaminade, T. et al. (2015). Anthropomorphic Bias Found in Typically Developing Children is Not Found in Children with Autistic Spectrum Disorder. Autism. 19(2): 248-251.
DeQuinzio, J.A, B. Taylor. (2015). Teaching Children with Autism Discriminate the Reinforced and Nonreinforced Res ponses of Others: Implications for Observational Learning. Journal of Applied Behavior Analysis, 48 (1):38-51.
Fauconnier, G. (1997). Mappings in thought and language. Cambridge: Cambridge University Press.
Fillmore, C.J. (1985). Frames and the semantics of understanding. Quaderni di Semantica, 1(2):222-254.
Ibanez, F.J.R.M. & Aransaez, C. (1998). Conceptual Schemas as propositional idealized cognitive models: in search of a unified framework for the analysis of knowledge organization. C.I.F. XXXIII-XXIV: 257-270.
Kim, E.S. et al. (2012) Bridging the Research Gap: Making HRI Useful to Individuals with Autism. Journal of Human-Robot Interaction, 1(1): 26-54. DOI 10.5898/JHRI.1.1.Kim
MacWhinney, B. (2000). The CHILDES PROJECT: Tools for Analyzing Talk (3rd ed.). Lawrence Erlbaum Associates.
Polzenhagen, F., & Xia, X. (2014). Language, Culture and Prototypicality. In F. Sharifian (Ed.), The Routledge Handbook of Language and Culture: 253–269.
Shamsuddin, S. et al. (2012). Humanoid Robot NAO Interacting with Autistic Children of Moderately Impaired Intelli gence to Augment Communication Skills. Procedia Engineering. 41: 1533-1538.
Sebeok, T.A. & Danesi, M. (2000). The Forms of Meaning: Modeling Systems Theory and Semiotics. Berlin/New York: Mouton de Gruyter.
Tager-Flusberg, H. (1985). The Conceptual Basis for Referential Word Meaning in Children with Autism. Child Develop ment. 56(5):1167-1178.
Tanaka, F., & Matsuzoe, S. (2012). Children Teach a Care-Receiving Robot to Promote Their Learning: Field Experiments in a Classroom for Vocabulary Learning. Journal of Human-Robot Interaction, 1(1): 78–95. doi:10.5898/JHRI.1.1.Tanaka