Nudges with conversational agents and social robots: a first experiment with children at a primary school

Hugues Ali Mehenni, Sofiya Kobylyanskaya, Ioana Vasilescu, Laurence Devillers

Nudges with conversational agents and social robots: a first experiment with children at a primary school

Hugues Ali Mehenni, Sofiya Kobylyanskaya, Ioana Vasilescu, Laurence Devillers

Abstract This paper presents an experimental protocol during which human interlocutors interact with a dialog system capable to nudge, i.e. to influence through indirect suggestions which can affect the behaviour and the decision making. This first experiment was undertaken upon a population of young children with ages ranging from 5 to 10 years. The experiment was built to acquire video and audio data highlighting the propensity to nudge of automatic agents, whether they are humanoid robots or conversational agents and to point out potential biases human interlocutors may have when conversing with them. Dialogues carried with three types of agents were compared: a conversational agent (Google Home adapted for the experiment), a social robot (Pepper from Softbank Robotics) and a human. 91 French speaking children participated in this first experiment which took place in a private primary school. Dialogues are manually orthographically transcribed and annotated in terms of mental states (emotion, understanding, interest, etc.), affect bursts and language register, which form altogether what we call a user state. We report on an automatic user states detection experiment based on paralinguistic cues in order to build a future automatic nudging system that adapts to the user. First results highlight that the conversational agent and the robot are more influential in nudging children than the human interlocutor.

Keywords: human-robot interaction, nudge, sensitive population, dialog system, paralinguistic features, filled and empty pauses

Hugues Ali Mehenni
CNRS, LIMSI, Paris-Saclay University
e-mail: alimehenni@limsi.fr

Sofiya Kobylyanskaya
e-mail: skobyl@limsi.fr

Ioana Vasilescu
e-mail: ioana@limsi.fr

Laurence Devillers
CNRS, LIMSI, Paris-Saclay University, Sorbonne University
e-mail: devil@limsi.fr

1 Introduction

The increasing usage of conversational robots in many everyday situations raises the question of their influence on humans [13]. However, this influence is hard to measure and often disregarded. This paper addresses the issue of the influence of the conversational agents on human users, and focus on "nudges" : indirect suggestions which can affect the behaviour and the decision making. The notion of "nudging" first came to light in 2008, proposed by R.H. Thaler (Nobel Prize in Behavioural Economy, Nov 2017) and C.R. Sunstein [15]. They stressed the fact that "nudging" was a tactic to subtly modify a person's behaviour, without restricting that person's choice. Indeed, nudging mainly operates through the affective system or by exploiting common cognitive bias (e.g. attention, memory, lazyness) [1].

Nudges could have a large impact on society, both negative and positive. On the one hand, they pose a threat to privacy [1] since people can be incited to leak their personal information. On the other hand, they could be used to improve efficiently and smoothly a vast number of tasks from diverse fields, e.g. education (attention, memory), transportation, health care. For now, "nudging" as a research topic has been covered mostly in behavioural economics [15].

The long term goal of this work is both to build an automatic dialog system able to nudge and to measure the influence of nudges exerted by conversational agents and robots on humans in order to raise the awareness of their use or misuse and open an ethical reflection on their consequences. The experiment was undertaken as part of the project "Bad Nudge Bad Robot" focusing on the modeling of nudging strategies within a spoken dialogue, funded by the DATAIA institute [2]. This project is part of the AI chair HUMAAINE: HUman-MAchine Affective Interaction Ethics (headed by L. Devillers) at CNRS and DATAIA. The chair team is composed of researchers in computer science, linguists and behavioral economists from Paris-Saclay University.

Here nudges are considered within a new research paradigm, the human vs conversational robots and/or agents verbal interaction. Data are collected trough a dialog protocol built to convey nudges. The aim of the experiment is to determine whether automatic learning allows measuring the influence of verbal nudges on a human engaged in a dialog with an automatic system.

2 Experimental framework and methodology

The experiment was conducted in a private primary school in early July 2019 as part of a multidisciplinary project involving researchers in Spoken Language Processing at the LIMSI-CNRS laboratory and in economy (RITM, Paris-Sud / Paris-Saclay University).

[1] https://medium.com/better-humans/cognitive-bias-cheat-sheet-55a472476b18

[2] https://dataia.eu/

The rationale behind this experiment is that human behaviour may be subject to cognitive bias when the interlocutor is a robot or a conversational agent. Indeed, humans tend to anthropomorphize machines [4] and to project emotions on them. This experiment also targets a sensitive population, children, who are more likely to be influenced. However, a preliminary experiment involving adults pointed out that they can also be influenced by social robots [5].

2.1 Experimental design

The design of the experiment consisted of a child interacting with a conversational partner for approximately 5-10 minutes. Children's age ranged from 5 to 10 (that is all the primary school levels). To observe the bias a child may display towards robots, volunteers were equally divided into 3 groups, corresponding to a balanced distribution in terms of age and gender. Each group was paired with a different conversational partner : (1) with a humanoid robot (Pepper from Softbank Robotics), (2) with a speaker (Google Home, adapted to the task) and (3) with a human (a PhD student participating in the project, aware of the experiment but working on a different topic than human-machine dialog).

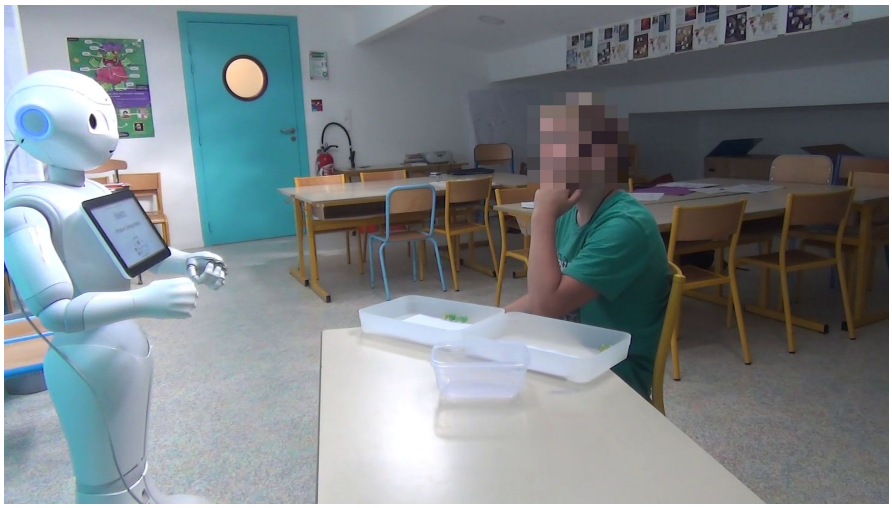

Fig. 1 Child volunteer engaged in a conversation with the robot Pepper

The experiment was conducted towards a Wizard of Oz (Woz) procedure. The dialogue was scripted and both the robot and the Google Home speaker were manipulated by a researcher during the interaction with the children, unbeknownst to the latter.

2.2 Structure of the experiment

The experiment is divided into 3 parts. The rationale behind this structure is to fits several analytic dimensions with respect to the various issues concerning nudges, that is behavioural economics, the propensity of a dialog system to implement nudging strategies and the dialog and emotional specificity of a population of children, known both as vulnerable and increasingly confronted to such systems. Prior to the experiment we requested and obtained the official approval of the ethical comity of Paris-Saclay University.

The three parts are as following:

1. Dictator Game, adapted to children. The game is well-known within the Game Theory [2] and is intended to measure altruism. In the present case, we give to a child a certain amount of marbles (for adults we would use money), and ask him to choose how much he wants to keep for himself and how much he is willing to give to another person (in the current configuration, another classmate, not specifically mentioned). We then try to influence the decision by using anchoring techniques (e.g. peer-effect strategies).
2. Open question, testing the amount of the confidence a child attributes to the discussion partner.
3. Quiz addressed to the children. The selected topic was video games and we investigated in advance that children would held the questions concerning some specific games. Specific dialog strategies (e.g. question repeated at several speaker turn distance) were employed during the quiz, to measure the attention and estimate which nudging strategies are the most efficient to raise children's trust and potentially affection towards a robot.

The experiment was followed by a discussion with all the children, during which we explained how the robot worked in an understandable way. The aim of this last action was to sensitize the young population about the potential misuse of nudges.

In the following sections we will describe the results of the experiments according to the 3 steps described above and propose a first detection system using machine learning techniques and paralinguistic cues associated to reactions (mental states, affect bursts, choice of language register) to nudging strategies. The section bellow focuses (Section 3) on a short description of the Dictator game experiment. Although this experiment implies few spoken parts, it provides a first overview of the children behavior as a function of the interlocutor employing nudging techniques. Section 4 focuses on the description of the corpus, the annotation strategy and the kappa calculation. Results are displayed in section 5 and concern both automatic detection of mental states in a broad sense (including also language register and affect bursts), based on paralinguistic features, and the linguistic description of the correlation between non-verbal information (here filled pauses) and annotation labels. We finally propose a discussion and further work objectives in section 6.

3 Towards quantifying nudging strategies: the Dictator game

The first part of our experiment, the Dictator game, enabled us to get concrete and quantifiable data on how much children were influenced by their interlocutor (cf Table 5). Although verbal data remains limited during this part of the experiment, the results provide insights on the patterns of interaction as a function of the type of interlocutor (robot, conversational agent, human). To start, children were invited to make a first decision on how they would like to distribute a fixed number of marbles. They were then subjected to 2 successive nudges using 2 different anchoring techniques (peer-effect and first-person strategies). We measured on average an altruism of about 45% during the experiment. About 50% of the children were influenced by their interlocutor and changed their initial choice during each nudging attempt. Furthermore, during the first and second nudging attempt, children were more influenced by non-human interlocutors (Google Home speaker or robot) than by the human one (cf Table 5). Indeed, among the children who changed their choice during the second nudge, 16% interacted with the adult whereas 40% and 44% respectively interacted with the robot and the Google Home speaker. This result is consistent with the hypothesis that a non-human interlocutor is more likely to influence. However, in order to generalize the observation, additional data would be necessary and should cover other groups of children as well as older volunteers (adults, elderly).

In order to quantify the nudging effects during the dictator game, a simple metric has been retained that consists of measuring how much (in terms of marbles) the child's choice got closer to the value of the nudge given by the interlocutor.

Nudge metric = (Difference in absolute value between the result before the nudge and the value given during the nudge) - (Difference in absolute value between the result after the nudge and the value given during the nudge)

Although, this metric gives an estimation of the number of marbles modified by a child and the direction of the change (positive or negative result of the metric), it does not show whether the child completely complied with the value given during the nudge or only approached it. To take this aspect into account, the last nudge metric was divided by the difference in absolute value between the result before the nudge and the value given during the nudge.

Normalized Nudge metric = Nudge metric / (Difference in absolute value between the result before the nudge and the value given during the nudge)

The metrics underlined that children are more influenced by the robot and the Google Home speaker than by the adult. During the game, different children behaviours can be observed : amusement, nervousness, doubts, etc. This led us to focus on these aspects in Section 5.

Table 1 Data collected during the Dictator game

Nb of children (age : 5-10)	Adult : 31	Robot : 29	Speaker : 31	Total : 91
Altruism Marbles (mean)	At first : 4.583	1st nudge : 4.7	2nd nudge : 4.427	Mean : 4.57
Nudged children 1st nudge	Adult : 15	Robot : 12	Speaker : 19	Total : 46
Nudged children 2nd nudge	Adult : 7	Robot : 17	Speaker : 19	Total : 43
Nudge metric 1st nudge	Adult : 0.68	Robot : 1.10	Speaker : 1.26	Total : 1.01
Normalized Nudge metric 1st nudge	Adult : 0.12	Robot : 0.30	Speaker : 0.26	Total : 0.23
Nudge metric 2nd nudge	Adult : 0.97	Robot : 1.90	Speaker : 1.48	Total : 1.44
Normalized Nudge metric 2nd nudge	Adult : 0.19	Robot : 0.43	Speaker : 0.46	Total : 0.35

4 Corpus

This section focuses on data description, annotation strategy and agreement between the annotators.

4.1 Data description

The experiment was conducted with 91 children, divided into 3 groups : one interacted with the social robot Pepper, one with the Google Home speaker and a third one with the human researcher. Table 2 bellow sumps up the distribution of the participants across the three configurations.

Table 2 Data description

Number of children	91
Number of male children	53
Number of female children	38
Number children with the Robot	29
Number children with the Speaker	31
Number children with the Human	31
Age of the children	5-10
Mean length of dialogues	8min10s
Total corpus duration	12h

4.2 Annotations

The corpus has been annotated with different labels at the speaker turn level in order to further correlate the characteristics of the user state at the turn and paralinguistic features. The labels are used to train classifiers with machine learning techniques in order to automatically detect the audio information (in this experiment, paralinguistic) that would help the dialogue system to better assess the state of the user. Plutchik's wheel [12] of emotions served as reference and the following six basic emotions were selected: amusement, respect, surprise, irritation, nervousness and intimidation (the 2 last labels describing two different low levels of fear). Furthermore, meta-labels opposed positive to negative emotions, in order to make the classification task easier for the models and to overcome limited data for some classes. "Attention" or "Interest" was also retained as label within the annotation system as such labels may help assessing whether a person is likely to be influenced or not. The label was associated to 2 classes : "Interest" and "Disinterest" (which could be linked to boredom, another emotion on Plutchik's wheel). Along the same lines, we then considered two additional annotation labels involving binary classes allowing to describe the mental state of the child: "Confidence"/"Doubt" and "Understanding"/"Confusion". Therefore, in this annotation system, mental states broadly designate emotion labels and labels relative to "Doubt", "Interest" and "Understanding".

We also annotated affect bursts : "Laughter", "Hesitation" and "Breath"; although, the number of chunks annotated with the label "Breath" are few and do not allow a machine learning classification.

We added two language register labels "Polite"/"Colloquial" as well, so as to further correlate the level of language children would adopt as a function of the interlocutor.

Finally, the label "Other" mainly contains neutral chunks and serves as a default class for the classifiers. Table 3 bellow sums up the annotation labels and the number of speech chunks per label.

Table 3 Description of the annotations and number of speech chunks per class (91 dialogs)

Annotations of mental states, affect bursts and language registers	**Number**
Positive Emotions (amused, surprised, respectful)	354
Negative Emotions (irritated, nervous, intimidated)	59
Confidence	622
Doubt	286
Interested	330
Disinterested	49
Understanding	68
Confusion	101
Polite	49
Colloquial	117
Hesitation	273
Laughter	52

The corpus data was annotated (cf figure 2) with the software ELAN [3] and the audio was extracted using Praat [4] scripts. The corpus also benefited from an orthographic manual transcription.

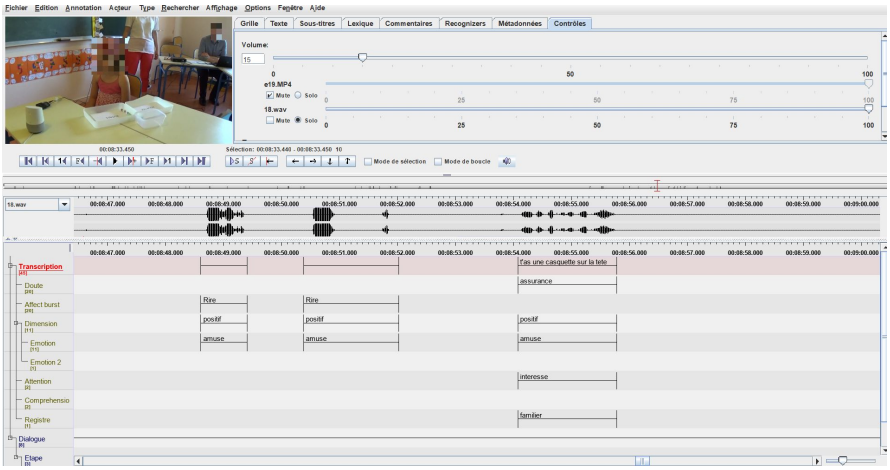

Fig. 2 Annotations of the user state (mental states, affect bursts, language register) with ELAN

4.3 Kappa : Inter-rater agreement

In order to measure the quality of the annotations, we use a control protocol: 5 files of each category (15 files in total) were annotated by two annotators then the Cohen's kappa coefficient was calculated. At this point, the emotion label contained also an "Indeterminate" class to be sure to take into account every potentially unclear emotions. We obtained 0.76 of agreement for the emotion and 0.68 for the doubt labels. Both metrics correspond to a substantial level of agreement. Divergences concern the number of segments annotated by each annotator. To compute the coefficients, we took into account only the segments which were annotated by both annotators. We considered that if one segment was annotated by one annotator but was left without annotation by the other, the emotion of this segment was not sufficiently well marked. So, these segments could not have a strong influence on the performance of the further automatic classifiers.

Few divergences were observed for opposite categories, for example for negative and positive dimensions. Most of the differences in annotations were observed for pairs such as indeterminate/negative, indeterminate/positive, intimidated/nervous

[3] https://tla.mpi.nl/tools/tla-tools/elan/

[4] https://praat.fr.softonic.com/

Table 4 Annotated segments

Annotated by A1 and A2 :	Annotated by A1 but not A2 :	Annotated by A2 but not A1 :	Total
134	205	73	412
0.33	0.49	0.18	1.0

which are more likely to be confused, and may strongly depend on the cultural and personal backgrounds of the annotators. In our future data collection, we will reproduce this control protocol several times during the annotation phase in order to improve the inter-annotator agreement.

Table 5 Annotation divergences

posit./negat.	posit./indet.	negat./indet.	intimidated/nervous
0.06	0.22	0.20	0.38

5 Results : Automatic detection of mental states, affect bursts and language register based on paralinguistic features and linguistic analysis of non-verbal information

This section focuses on the automatic detection of the user state based on paralinguistic features and on the linguistic analysis of this non-verbal information. Recall that the main focus of the experiment is to automatically detect and classify mental states (emotion, Understanding, Doubt, Interest), affect bursts and language registers through paralinguistic information. The scores obtained from the automatic classifiers are presented in the section 5.1, whereas the section 5.2 focuses on the linguistic description of the correlation between non-verbal information (here filled pauses) and mental states.

5.1 Paralinguistic detection of the user state

This section focuses on the detection of mental states, affect bursts and language register based on paralinguistic information as part of a future automatic dialogue system (which is further described Section 6). Paralinguistic studies about emotion with SVM (Support Vector Machine) or other conventional approaches have been conducted in previous works and a specific interest was given to minimalistic acoustic parameter sets, e.g. GeMAPS (the Geneva minimalistic acoustic parameter set) for voice research and affective computing [8] or robust small sets [14]. Here par-

alinguistic parameter sets were also used to implement classifiers with speech data specific to nudges.

We used the software OpenSMILE [5] to extract relevant features in order to train the models (emobase2010 features of OpenSMILE [9]). The results are provided in Table 6. The models implemented were SVM (Support Vector Machine) and Random Forest.

Each classifier had 2 or 3 classes to classify. The results given for each set of classes are the mean F1-scores (with standard deviation), which evaluate how well a classifier discriminates one class from another. The scores were computed with nested cross validation and the data for each class was weighted in the classification models so as to take into account the data imbalance.

Table 6 Mean F1-scores for mental states, affect bursts and language registers classifiers (10-fold nested cross validation) with the emobase2010 features (1581 feat.)

Classes	SVM	R.F.
Positive emotions / Negative emo. / Other	**0.56** ±0.01	0.45 ±0.03
Interested / Disinterested / Other	**0.56** ± 0.01	0.49 ± 0.01
Doubt / Confidence / Other	**0.59** ±0.01	0.48 ±0.01
Understanding / Not Understanding	0.65 ±0.04	**0.66** ±0.02
Polite / Colloquial	**0.82** ±0.01	0.79 ±0.03
Hesitation / Laughter / Other	**0.71** ±0.01	0.64 ±0.01

SVM classifiers seem to produce better results than Random Forest classifiers with our dataset.

These results show that some pairs/triplets of ”states” are more easily distinguishable than others on the basis of the paralinguistic features. Positive and negative emotions are indeed harder to classify than affect bursts such as hesitation and laughter. The small size of the dataset may explain the relatively high variances of some scores and the classifiers would probably benefit from more data. Another constraint the classifiers had to cope with was the noise in some speech chunks because of the noisy environment of the school. This brings some robustness to the models and more realistic scores.

Given the performance of the classifiers, they will be used as a weighted input for the dialogue manager rather than as a discrete input. A vector gathering all of their predictions in real time will then give relevant information about the state of the user to a reinforcement algorithm in the dialogue manager (further discussed section 6).

[5] https://www.audeering.com/opensmile/

5.2 Contribution of non-verbal information for mental states characterisation

This section focuses on the role of non-verbal information such as filled (e.g. in French “euh” and “bah”) and silent pauses in mental states characterization. 15 dialogues are considered for the analysis (5 dialogues for each pair infant vs robot/Google Home speaker/human) in order to estimate if the non-verbal information can provide reliable cues about the speaker’s state which can further be correlated to the prediction of the classifier. Filled pauses are key elements in dialog construction and management [3] [7]. They carry also salient paralinguistic information and can be language dependent. Besides, it has been shown that the position of a vocalic hesitation within the speaker turn can be correlated with various functions such as keeping the floor, introducing new information or manifesting the intention to close the dialog [16]. Here the filled pauses are considered in three positions within the speaker turn, that is initial, internal and final. As for the silent pauses, we consider so far the speaker turn internal ones. Indeed, the pauses observed at the beginning of a speaker turn can be decoded as moments of latency between the question of the human, robot or Google Home speaker and the child’s response and consequently, were analysed separately. The occurrence of filled (hesitations) and empty pauses is considered as functions of different pairs of mental states, that are Doubt/Confidence, Positive/Negative emotions, Intimidated/Nervous negative emotional state, Interested/Disinterested (Table 7). Although previous research pointed out the correlation between hesitations and negative or at least non-neutral mental states [6], the current analysis does not point out a strong correlation between pauses and non-neutral states such as doubt.

The number of hesitations and silent pauses corresponding to speech chunks labeled as ”Confidence”, ”Positive emotions”, ”Intimidated” and ”Interested” is superior to the remaining labels. This observation applies for all the three groups (human, robot, Google Home speaker). The label ”Intimidated” which roughly corresponds to a discomfort or stressful attitude of the children in front of the agent, and the quantity of hesitations and pauses are correlated. In the observed examples, if the child seems intimidated, the amount of pauses and hesitations produced increases, however the label ”Nervous” does not seem to elicit an increasing number of such vocal items.

Finally, the mean duration of pauses and hesitations (Table 8) corresponding to the speech excerpts labeled ”Doubt” are in most of the cases superior to those labeled ”Confidence”, despite the fact that they are more frequently observed. This observation is true for the interactions with a human and a robot.

The preliminary results above point out the potential correlation between the incidence of non-verbal information and user states and it is promising for further integration within the automatic detection experiments.

Table 7 Number of pauses and hesitations introducing mental states

	human		robot		speaker	
	hesit.	pauses	hesit.	pauses	hesit.	pauses
conf.	46%	50%	23%	23%	50%	60%
doubt	17%	36%	19%	13%	16%	16%
pos.	23%	-	10 %	3%	36%	32%
neg.	-	-	-	3%	-	-
intimid.	20%	-	6%	36%	7%	20%
nervous	-	-	-	-	-	-
interest.	43%	36%	26%	44%	41%	32%
disinter.	-	-	10%	5%	2%	4%

Table 8 Mean duration of hesitations and pauses (in s)

	human		robot		speaker	
	hesit.	pauses	hesit.	pauses	hesit.	pauses
conf.	0.86	10.41	0.72	1.53	0.85	2.13
doubt	1.11	13.62	0.80	3.28	1.16	1.78

6 Towards an automated nudging dialogue system

The next step of this project is then to build an automated dialogue system. The user state detection (described in Section 5.1) will thus be used in the Spoken Language Understanding (SLU) part of the dialogue system. It will be built as modular parts and will be coupled with semantics for a better understanding. The different classifiers trained and explained in Section 5.1 will be fed with speech chunks during new conversations and produce in real time a probabilistic distribution for each set of classes. These predictions will be then gathered in a vector, which will be given as input for the dialogue manager.

For the dialogue manager model, we are working on a POMDP-based architecture [17] with online learning. Reinforcement learning algorithms will be used in order for our agent to adapt to its interlocutor and learn the most efficient nudging strategies for him. We will thus use a sample-efficient algorithm such as the Kalman Temporal Differences model (KTD)[11], whose advantages were explained for instance in E. Ferreira's thesis[10].

7 Conclusion

The study presented in this paper is part of a larger project whose aim is to measure the influence of nudges exerted by conversational agents and robots on humans, in order to raise the awareness of their use or misuse and to open an ethical reflection on the consequences. Two underlying objectives are to design an automatic dialog system able to nudge and to evaluate its feasibility in realistic conditions. The present paper focuses on these 2 objectives: we describe a preliminary Wizard of

Oz experiment built for collecting a corpus of dialogues in real-life situation and a detection system based on paralinguistic features. The corpus consists of dialogues between three types of interlocutors (the robot Pepper, a Google Home speaker used as a conversational agent and a human) and children, recorded in a primary school. The rationale behind this choice is that children are a sensitive population and that they are assigned to increasingly interact which such tools. 91 children with ages ranging from 5 to 10 participated in the experiment. Firstly, a Dictator game is proposed to the volunteers under the form of a marble game. To quantify the nudging effects during the Dictator game, a simple metric is retained that consists of measuring the difference between the value of the nudge proposed by the interlocutor (as amount of marbles) and the choice of the children (the effective amount of marbles selected by the children). Thereafter, the corpus collected is annotated in terms of emotions, attitudes, affect bursts, disfluences and language register in order to provide an input for the detection system aimed to automatically identify mental states in a broad sense, correlated to nudging strategies. Moreover, the contribution of the non-verbal information (filled ans silent pauses) for user states characterisation is estimated and seems a promising lead for improving the automatic detection. The first results highlight that the conversational agent and the robot are more influential in nudging children than the human interlocutor.

The next step of this project is to build an automated dialogue system able to nudge. Future work is also focusing on collecting additional data from both children and other sensitive (elderly) and non sensitive populations. The results will feed into a more general reflection on the nudges and the ethical issues they raise within the activities of the AI research chair HUMAAINE.

Acknowledgements The study was funded by the DATAIA project "Bad Nudge, Bad Robot".

References

1. Acquisti A. et al., (2017), Nudges for Privacy and Security: Understanding and Assisting Users' Choices Online, ACM Computing Surveys, Vol. 50, No. 3, Article 44, August.
2. Benenson J.F., Pascoe J., Radmore N., (2007), "Children's altruistic behavior in the dictator game", Evolution and Human Behavior.
3. Clark H.H., Fox Tree J.E., (2002), "Using uh and um in spontaneous speaking", Cognition, 84, 73–111.
4. Reeves B., Nass C., (1996), "The media equation".
5. Devillers L., et al., (2015), "Inference of human beings' emotional states from speech in human–robot interactions", International Journal of Social Robotics 7 (4), 451-463.
6. Devillers L., Vasilescu I., Vidrascu L., (2004), "Anger versus Fear detection in recorded conversations", Proceedings of speech prosody.
7. Duez D., (2001), "Signification des hésitations dans la production et la perception de la parole spontanée", Parole 17/19, 113-137.
8. Eyben F., Scherer R., Schuller B., Sundberg J., André E., Busso C., Devillers L., Epps J., Laukka P., Narayanan S., Truong K., (2015), "The Geneva minimalistic acoustic parameter set (GeMAPS) for voice research and affective computing", IEEE transactions on affective computing, Volume 7, Num 2,Pages 190-202, editor IEEE

9. Eyben F., Weninger F., Gross F., Schuller B.,(October 2013), “Recent Developments in openSMILE, the Munich Open-Source Multimedia Feature Extractor”, In Proc. ACM Multimedia (MM), Barcelona, Spain, ACM, ISBN 978-1-4503-2404-5, pp. 835-838, doi:10.1145/2502081.2502224
10. Ferreira E., (2015), Apprentissage automatique en ligne pour un dialogue homme-machine situé, Université d’Avignon.
11. Geist M., Pietquin O., (2010), ”Kalman temporal differences”, Artificial Intelligence Research 39(1), 483–532.
12. Plutchik R., (1980), ”A general psychoevolutionary theory of emotion”, Emotion : Theory, research, and experience, Vol. 1, Theories of Emotion, (p. 3-33), New York: Academic.
13. Sciuto A., Saini A., Forlizzi J., Hong J.I., (2018), “Hey Alexa, What’s Up?”, Proceedings of the 2018 on Designing Interactive Systems Conference 2018 - DIS ’18. doi:10.1145/3196709.3196772
14. Tahon M., Devillers L., (2015), ”Towards a small set of robust acoustic features for emotion recognition: challenges”, IEEE/ACM transactions on audio, speech, and language processing, volume 24, Num 1, P16-28, Ed. IEEE
15. Thaler R. H., Sunstein C. R., (2008), ”Nudge: Improving decisions about health, wealth, and happiness”, New Haven, CT : Yale University Press.
16. Vasilescu I., Rosset S., Adda-Decker M., (2010), ”On the Role of Discourse Markers in Interactive Spoken Question Answering Systems”, Proceedings of the Seventh International Conference on Language Resources and Evaluation.
17. Young S. et al., (2013), ”POMDP-based Statistical Spoken Dialogue Systems : a Review”, IEEE.