GOVERNMENT OF THE RUSSIAN FEDERATION

Federal state autonomous educational

institution of higher education

National Research University

Higher School of Economics

Department of Humanities

School of Fundamental and

Computational Linguistics

Tsfasman Maria

DEVELOPING EYE GAZE SYSTEM TO SUPPORT HUMAN-DIRECTED COMMUNICATION IN F-2 EMOTIONAL ROBOT

Bachelor dissertation of 4th year Bachelor's student, 142 group

Academic head of the educational program	Scientific supervisor
Cand. philological sciences, Assoc.	Cand. philological sciences, Assoc.
Yu.A. Lander	A.A. Kotov
___________________________	___________________________
« » __________ 2018 г.	« » __________ 2018 г.

Moscow 2018

0. Abstract

When approaching eye gaze simulation in conversational agents developers tend to focus on maximum resemblance to human-like oculomotor apparatus. However, is that crucial when constructing an artificial conversational partner? Cartoon characters do not have human-like saccades or human-like eyeball structure - nevertheless, we tend to develop a greater emotional attachment to them than to a human stranger. The focus of F-2 Emotional Robot project is to create a conversational agent which may not resemble a human in its physiology but which could communicate through both verbal and body languages. This paper aims to study and model the eye gaze moderation system aimed at making F-2 Emotional Robot more attractive and natural in human-robot interaction. This project proceeds from the view of oculomotor behavior as an information-transmitting system, consisting of intentional and spontaneous signs in communication. First, I present my analysis and description of human oculomotor conversational behavior, a system of communicationally meaningful expressions, on the basis of the Russian Emotional Corpus. Next, I implement the suggested model in the robot. I prove the significance of eye gaze in human-robot interaction through a short evaluation experiment. As a final step of my research, I assess perception of resulting eye gaze moderating system by means of user-experience testing within human-robot interaction. The resulting eye gaze model is not exclusively tailored to F-2 Emotional Robots, it can be used as a theoretical basis for constructing eye gaze behavior in conversational interfaces in general. My research presents the practical implementation of eye gaze behavior in F-2 Emotional Robot, as well as contributes to the broader field of study of the non-verbal communication systems.

1. Introduction

Background. Oculomotor behavior has shown to be an important component in human interaction. It is a complex non-verbal system, transmitting information through intentional and spontaneous signs.

Eye movement simulation is a well-described feature in robotics, which is implemented in many robots and conversational interfaces. The body of research on human oculomotor behavior and its design in communicational agents can be divided into the two major areas: human-oriented (cognitive, psychological, and linguistic) studies, and projects with a technical bias (involving AI development, medicine, and physics). The former researchers focus on the meaning and psychology of eye gaze regardless of its quantitative characteristics, such as the timing of eye movement and the angle of direction. The latter primarily consider the anatomical and physiological properties of oculomotor behavior, focusing on its quantitative description. My research aims to combine the approaches from two mentioned perspectives.

Problem statement. Developers of robot eye movement usually strive to achieve a maximum resemblance to the human eye gaze apparatus, with a focus on human-like saccades, adjustable pupil size and the anatomical oculomotor structure; thus they do not consider the communicational significance of eye movements. This approach leaves aside the question whether human-like appearance is the main trigger of establishing an emotional bonding between humans and humanoid creatures, or are there other factors involved. Evidence suggests that we tend to form a stronger attachment to dogs and cartoon characters rather than to objects that exhibit human-like resemblance (Mori, 1970). In fact, human history proves that being generally similar to us but differing in a few minor aspects can is likely to become a cause of a conflict rather than a source of empathy (Yuval Noah Harari, 2011). Thus, as well as the main objective of the F-2 Emotional Robot project in general, I aim to create a system generating meaningful eye gaze behavior rather than imitating any human physiological features.

Delimitations of the study. My work is conducted as a part of the F-2 Emotional Robot project developed by the Laboratory of Cognitive Technologies at Kurchatov Research Institute in Moscow, Russia (Kotov & Budyanskaya, 2012). The practical challenge of designing meaningful eye gaze behavior in the robot is the major objective in my project; however, its secondary aim is much broader: to investigate, annotate and describe human eye movement behavior in a dialogue. The resulting eye gaze system is to increase robot's believability and likability in human-robot interaction.

I proceed from the view of oculomotor behavior as a non-verbal conversational system and classify the strategies of meaningful eye movements, the contexts of their usage, their specifics and the patterns of their interaction. After that, I test the validity of my model in human-robot interaction and identify the strategies that are more attractive and realistic from users' perspective.

Robot F-2's gesture system is based on microstates system generating the macrodynamics of the whole system. Microstates are gesture packages, determining a set of robot's head, eye, hands, and body movements. Certain internal and external stimulus send those microstates to robot's behavioral queue. Every microstate sent to the queue has its weight determined. The weight signifies how important the package is for the robot to complete. If microstate's weight is higher than that of states in the queue, robot stops gesticulating the microstates in the queue in order to complete the one with higher weight; after that robot returns to the queue and continues gesticulating interrupted state, unless it loses its weight by that time. My aim was to construct a list of eye movement microstates and a system that determines their activation depending on different internal and external conversational features. The resulting system was supposed to moderate meaningful oculomotor behavior of the robot in conversation, making robot more attractive and convincing.

Research questions. In my project I discover and classify significant non-verbal eye gaze signs as well as prove their significance by designing and performing them on robot's eye movement system.

The research questions this paper seeks to answer are following:

- What symbols the system of oculomotor communication consists of? How do those symbols interact between each other? Which contexts are different eye gaze symbols used in and what do they mean?
- Is this system physiological only or does it transmit certain information about an interlocutor? How relevant is robot's eye gaze in human-robot interaction?
- How can meaningful eye movements be designed on the robot? How should eye gaze strategies interact within the robot's system? How to structure eye gaze communication model so that it moderates adequate eye gaze behavior automatically?
- How does my model affect people's perception of the robot?

Methods and tools. My research is divided in four parts and is based on multimodal corpus analysis; different tools meant to design gaze features and their interaction in the robot; and user experience testing, including semantic differential scaling tests and other evaluation tasks.

Oculomotor behavior classification is based on the Russian Emotional Corpus, which stores video data of dialogues transcribed by face expressions: head, mouth, eyebrows, and eye movements. Basing on the statistics from the corpus, I discover the most common non-verbal eye expressions and the context they are usually used in (Section 3). Then, I find whether eye gaze is noticeable and relevant for users, when listening to robot, by means of preliminary evaluation experiment (Section 4). After discovering different eye gaze strategies I design them as a software library for the robot (Section 5). The algorithms are then tested on users in order to discover how significant certain gaze signs are in a dialogue (Section 6).

Thus, in a general attempt to design believable eye movement for the robot, my research is to have four intermediary results: a corpus-based model of human oculomotor conversational behavior; a verification of eye gaze significance in human-robot interaction; eye gaze moderating system designed on F-2 Emotional Robot; and the proof of the model significance, based on a human-robot interaction experiment.

2. Literature Review

In this section I will describe studies conducted in the area of human oculomotor behavior and its simulation on robots and agents.

Why are specific eye movements that essential in human interaction? Eye gaze communication salience appears to be adaptive by its nature and often occurs in different mammal species as an important communication system (Baron-Cohen, 1995). Ability to recognise and pay particular attention to eye region is innate: 14-26 weeks old children are already able to differentiate between avert and direct gaze (Hains & Muir, 1996). Eye gaze strategies and perception has shown to correlate with several individual features, such as age, culture, and gender (Matsumoto, Hwang, & Frank, 2016).

When trying to model robot's eye gaze behavior the developer bumps into a serious problem. Within a great variety of eye behavior research papers, there are almost no works focusing on exact time spent looking at one or another point, exact pupil size, or exact gaze directions. Thus, this research had to, first, research those numbers through multimodal corpora. Nevertheless, there is a great variety of studies reporting the factors different oculomotor behavior depends on. Hence, in 2.1 I overview key works on eye movements from psychological, biological and, most relevantly, non-verbal linguistic perspectives.

Section 2.2, on the other hand, provides an overview of existing eye gaze simulations in conversational agents.

2.1 Eye movements in human interaction (theoretical perspective)

What makes us use a specific gesture in conversation? Scherer in his paper (1988) drew distinction between 2 types of factors causing body expressions: 'push factors, which determine affect expression mainly via physiological effects, and pull factors, which influence expression via socially mediated models' (Scherer 1988, p. 79). Although that terminology is essential to my research and will be used many times further, it is important to note that most of the times I do not distinguish those two types of factors. I will divide eye gaze triggers in three types: social, emotional, and linguistic. Social type (described in 2.1.1) includes eye gaze factors adaptive by their nature and typical for different mammal species. Emotional ones (described in 2.1.2) are the factors of eye gaze that are due to express or perceive emotions. The third type, linguistic, include eye gaze triggers and strategies that depend on linguistic features of the verbal information transferred or perceived while communicating.

That division is purely symbolic and serves organisational purposes more than represents diverse cognitive processes underlying body language. In humans, in fact, those factors are usually mixed together and hardly distinguished one from another. Nevertheless, in robot it was constructed as a set of competing factors triggering one or another oculomotor expression.

2.1.1 Social triggers

Social eye gaze functions described in the literature are **attraction**, **dominance**, and **conversational engagement**.

Such eye gaze role as getting attracted after reading certain eye gaze information appears to be one of the features that we inherited from our ancestors. It was important for the procreation to choose a partner which is more likely to accept. Pupil size and gaze direction plays a crucial role in showing and perceiving interest in the conversation. Men tend to be more attracted to women or men when the target has dilated pupils (Tombs & Silver 2004, Rieger & Savin-Williams, 2012). Regarding eye direction, Mason, Tatkow & Macrae (2005) discovered that when looking towards us, people appear to be more attractive and likable than when looking away (Mason, Tatkow & Macrae, 2005).

The second social function is eye gaze as a sign of **dominance or submission**. Tang & Schmeichel in their experiment (2015) provide an evidence that a long direct eye gaze is considered as a sign of dominance, as opposed to avert gaze - a sign of submission (Tang & Schmeichel, 2015). Thus, in certain cultures, for instance, Japanese, looking away is considered more polite than looking directly in another person's eyes (McCarthy, Lee, Itakura, & Muir, 2006).

As dominance is mostly expressed by the duration of direct eye gaze here it is important to note what duration is considered neutral in European culture. Different experiments got different results on that point. Gullberg and Holmqvist (2006) conclude that listener spends 90-95% looking at the speaker. However, further research has shown that in a less artificial environment it measures at around 40-45% towards the speaker (Grishina, 2017). The rates vary widely for the speaker's oculomotor behavior, depending on the topic, the style of the conversation/ monologue, and speaker's individual features.

The third and the last social eye gaze function is expressing a person's **conversational engagement**. While conversational disengagement can be expressed by looking away, a participant's interest and engagement in the dialogue is has a more complicated algorithm. Blaison et al. (2012) and Ishii, Ryo & Yukiko (2010) showed that the more a person mimicry interlocutor face expressions, including eye gaze direction, the more is a person engaged in their conversation.

2.1.2 Emotional triggers

Eye gaze expressions connected to emotional states are harder to classify and are not that well described in previous research. The reason is that emotions are a spectrum rather than a fixed set of options.

Here, the distinction between pull and push factors (Scherer, 1988) is specifically important. Pull-caused eye gaze expressions are the ones triggered by hormonal change caused by strong emotional reaction, for instance, widened eyes when experiencing shock. Push-caused, on the other hand are expressions illustrating speech - e.g. widened eyes when telling about how shocked you felt when you saw something.

2.1.3 Linguistic triggers

The last type of eye gaze factors is linguistic - connected to different speech features. It is thoroughly described in Grishina (2017) and Altmann (2011). Grishina's book on body language from a linguistic point of view (2017) presents one of the most detailed research on the issue. She signifies proves the importance of such factors as conversational role (listener or speaker), phrasal stress, theme and rheme, emphasis, negation and confirmation.

In conclusion, eye gaze triggers is one of the most researched questions in human oculomotor behavior. In my research I will try to use all the factors I mentioned above in order to include them in the robot eye gaze behavior system. However, what was left untouched in the previous research was how perceptually important those factors and proceeding eye gaze expressions were. Moreover, the studies usually focus one one particular type of factors, whereas I will try to embrace the most perceptually notable factors from those mentioned above - combining social, emotional, and linguistic triggers.

2.2 Robot eye gaze systems tradition

Virtual agent developers have been modelling human eye and its behavior models for a few decades already (Ruhland et al., 2015). Most of the models focus on anatomical resemblance of human oculomotor physiology and behavior. Francois et al. (2009), for example, designed their eye model based on iris photographs. Berard et al. (2014) also included sclera, and cornea anatomy in their eye design. Many developers also include such physiological detail as saccades, vestibulo-ocular reflex, eyelid movement (such as blinks) animation in their virtual eye models (Lee et al., 2002; Kolmogortsev et al., 2013; Gu et al., 2007; Sean et al., 2012; Admoni et al., 2013; Moubayed et al., 2012; etc.).

The question is whether physiological resemblance is essential when constructing a conversational agent? Are those details conversationally important in human-robot interaction? The assumption F-2 Emotional Robot project aims to prove is that a creature that shows empathy by means of verbal and body language is more trustworthy than somebody anatomically resembling a human (Kotov & Budyanskaya, 2012).

Yet, this project is not the first to use human-robot interaction as a way to determine human perception of body language. Ishii and Yukiko (2010) studied the perception of conversational engagement in human-agent communication. Admoni et al. (2013) researched how different visual attention patterns affect user perception of the robot. Bee, André & Tober (2009) built an embodied conversational agent able to flirt. The authors conclude that participants felt more attached to the flirting robot than to the one with neutral behavior. Komatsu & Takahashi (2013) proved the importance of joint attention while playing a game with a conversational agent. It appeared that eye contact does not obligatorily means emotional attachment to the robot, but eye contact in combination with joint attention correlate with better connection to the robot. However, those studies refer to a limited set of eye gaze symbols rather than to functioning conversational behavior. Nevertheless, they all confirm the significance of eye gaze in robot-human interaction.

To conclude, although oculomotor behavior in robots and humans has already been thoroughly studied and modelled, researchers and developers do not usually view eye gaze behavior as an information-transmitting complex system. That is why, in my project I discover the question of modelling a system controlling robot's oculomotor behavior.

My research is divided into four steps: discovering and classifying eye movements as meaningful non-verbal signs based on the multimodal corpus (Section 3); verifying whether eye gaze is important in human-robot interaction through an evaluation experiment(Section 4); designing those strategies on the robot itself (Section 5); proving the significance of my system in action by conducting a user experience experiment (Section 6).

3. Corpus analysis

As stated above, conversational eye gaze behavior is not described enough in literature to construct it on the robot without any additional preliminary research. Specifically, most studies of conversationally significant eye movements do not focus on concrete measurements of eye gaze duration, exact direction, and timing. Thus, the first step in my research was to determine significant eye gaze symbols and describe human oculomotor behavior in communication. My study is based on qualitative and quantitative corpus-analysis.

3.1 Corpus characteristics

There are quite a few Russian multimodal corpora. Most of them consist of either films, theatre plays, and TV-programs, e.g. the Multimedia Russian Corpus MuRCo (Grishina, 2009); or video interviews in an experimental environment, e.g. the Russian Multichannel Discourse Corpus (Kibrik et. al., 2018). Although that kind of corpora can be extremely efficient when studying certain topics, it is not particularly helpful for studying natural gestures and eye expressions.

In contrast, the Russian Emotional Corpus (Kotov & Budyanskaya, 2012) consists of video recordings of natural dialogues and interviews. The Russian Emotional Corpus (further REC) consists of three types of video recordings: students, passing exams; people appealing to communal services; and tea conversations with creative people about their career fields. Those situations are chosen in order to stimulate emotionality. The gathered material allows one to investigate spontaneous gestures and conversational behavior in emotional situations. All the participants are Russian native speakers of 19 - 35 years old.

The materials analysed in this paper included 127 minutes of speech video-recordings recorded from two informants. Two video-recordings include ‘rec-art01’ and ‘rec-art02’, I will refer to them by these names further.

3.2 Data annotation

REC videos are annotated by speech and body gestures, using ELAN. Video annotations include different tiers depending on what body part is transcribed: head, eyes, mouth, or hands. There is also a tier with speech transcription, divided in periods by syntagmatic borders.

Although corpus annotation is quite thorough in that it includes many different tags, oculomotor movement annotation consisted of five tags only, namely ‘side gaze’, ‘upward gaze’, ‘eyes closed’, ‘eyes narrowed’, ‘winking’. As this set of tags was not enough for my research, I had to introduce a new set of tags.

I examined the following three eye gaze features: direction, duration, and meaning. There is one more aspect that my annotation did not cover - eye expansion. Although that eye gaze feature is relevant for eye gaze perception, robot F-2 does not have agile eyelids yet. Although eye expansion aspect is still to be researched and designed in detail in further studies, I include that aspect in my description of eye gaze expressions where relevant.

Since ELAN annotation implies using a timeline and divides a recording into time periods, eye gaze duration did not require any special annotation. I created three separate tiers according to two other eye gaze features: ‘eye direction’ (3.3), and ‘gaze function'(3.4). In total, 1875 annotations were made on eye direction, and 1415 - on eye gaze functions.

3.3 Eye gaze direction annotation & distribution

Regarding direction, my annotation contains seven tags required to simulate gaze behavior on the robot: ‘right’, ‘left’, ‘upward’, ‘downward’, ‘towards the interlocutor’, ‘towards an object’, ‘closed’. The tier was formatted in such a way that every time period had a tag assigned. This method allows counting rates and applying statistics on the data.

The distribution of eye direction appeared to be relatively the same for both informants. Significantly different parameters are highlighted with red in Tables 1 and 2:

Table 1. *Eye direction distribution - rec-art01 informant*

rec-art01	**up**	**left**	**down**	**right**	**closed**	**object**	**eye contact**
% from total duration	4%	5%	18%	27%	1%	4%	41%
total number of annotations	26	29	111	170	9	27	255
median (sec)	1.27	1.24	1.85	1.36	1.48	1.56	3.93
min (sec)	0.25	0.32	0.21	0.29	0.31	0.47	0.42
max (sec)	7.59	4.66	8.23	11.73	3.83	2.94	38.58

Table 2. *Eye direction distribution - rec-art02 informant*

rec-art02	**up**	**left**	**down**	**right**	**closed**	**object**	**eye contact**
% from total duration	14%	27%	8%	10%	3%	3%	35%
total number of annotations	171	332	104	126	35	34	433
median	1.75	1.87	2.46	1.75	0.71	2.14	3.34
min	0.02	0.17	0.16	0.24	0.17	0.44	0.24
max	13.51	19.63	16.26	12.85	2.21	66.76	32.16

It appears that one informant looks to the left more than to the right (Table 1), and another informant - the other way around (Table 2). In my system, this fact is used as a slight argument to the randomised distribution of choice of the side when looking away form the speaker in some situations. What is more, my data is yet another proof of Grishina's (2017) results about the rate of eye contact in a neutral conversation, as opposed to Gullberg and Holmqvist's work (2006).

As an another solution to the problem of eye gaze design on the robot, one could moderate a randomized system with different probabilities of eye gaze direction. Probabilities could be taken from the resulting rates of eye gaze direction distribution. However, it contradicts the main assumption of my project: eye gaze behavior as an important information-transfering system in communication. That is why the next tier to be described consists of eye gaze meanings and is called 'gaze function'.

3.4 Eye gaze function annotation & description

The 'gaze function' tier had a more complicated annotation algorithm as I first had to determine the range of tags needed. 'Gaze function' describes the meaning of eye gaze; in other words, it resembles the explanation of any change in eye direction. The first 40 minutes of annotation were tagged in a freestyle way containing comments on eye gaze meaning. The resulting list of 496 comments consisted of functions that could be joined up in general groups but still had various subtypes. According to the resulting groups the set of 10 tags was introduced. The remaining 73 minutes of data were annotated using that system.

However, as the list was not enough to cover certain details about subgroups, an extra tier dependent on 'gaze function' tier served for comments and any detail about function's subgroups.

Table 3 below describes the whole system of tags and there subtypes, and their context. Table 3 presents not only tags, but is a resulting 'dictionary' with eye gaze expression classification, their definitions and context.

Table 3. *Eye gaze function 'dictionary'*

Function	Function subtype	Detail & explanation
thinking	***function itself***	'Thinking' function includes looking around as if a person is looking for an actual object in communication space.
	argumentation	The function is stronger activated when a person is 'looking' for arguments.
	on the answer	The function is specifically strong when a person is getting ready to answer to a question.
	remembering a word/a situation	Sometimes a person is 'looking' for a word or describing/remembering a situation.
	mental arithmetic	When counting something, people usually look up/ around them.
speaking		Normally, when telling a sentence a speaker looks interlocutor in the eye.
phrasal stress		Phrasal stress is frequently highlighted by eye contact. The reason can be that phrasal stress usually marks the most informative word in the sentence. Thus, by looking in the interlocutor's eye a person checks whether they understand the transferred information (connected to 'attention to person's reaction' function below).
enumeration		Sometimes when listing something, a person changes eye gaze direction, as if they were locating different objects in the communication space. This phenomenon appears in hand gesticulation more frequently (Grishina, 2012).

joking		Speaker's behavior is more complicated when joking. One strategy is to look at the interlocutor the whole joke, and then glance away. The second strategy is to look away during the whole joke and then look back to the interlocutor to check the reaction.
listening		When listening, a person usually looks at the speaker most of the time as a sign of conversational engagement. However, sometimes a person shows that he is thinking speaker's words through (see 'thinking' function).
attention to person's reaction	***function itself***	In order to show conversational engagement and to be social, speaker checks out interlocutor's reaction seeking to find it in their eyes/ face expression (see also 'speaking' and 'phrasal stress' functions above).
	whether an interlocutor understands	The function is particularly strong when a person saying something new/ complicated verifying whether something needs more explanation.
	question	After asking a question a person almost always looks at the speaker to show their interest in the answer.
mimicking person's eye gaze		As described above (2.1.1), mimicking interlocutor's face expression has shown to be an important sign of one's conversational engagement. My data shows several cases of the same phenomenon considering eye gaze direction - it usually occurs in the cases of joint attention to an object, but sometimes.
softening anti-social situation	***function itself***	'Softening anti-social situation' is one of the social eye gaze functions (see Section 2.1.1). In order to avoid being aggressive/ dominant a person looks aside when saying something socially inappropriate or when contradicting with other person's viewpoint.
	disagreeing	The function is particularly activating when disagreeing with/ contradicting something.
	softening 'strong' words	When swearing, or saying something anti-social/ hyper-emotional people tend to avoid eye contact.
	too long eye contact	That function is also activated when the eye contact was too long in order to avoid seeming dominant.

eye gaze pattern (iconic eye gaze)	***function itself***	‘Eye gaze pattern’ combines in itself specific eye gaze gestures. Eye gaze tends to be very metaphorical. It serves speech illustration not less than hands do. ‘Eye gaze pattern’ function, in fact, is more language specific than any other ones, as it is frequently connected to linguistic metaphors. It would be interesting to study those in multilingual context.
	verbs of (immediate) movement	Sometimes, when explicating a quick movement people illustrate it with a quick glance away, sometimes it includes corresponding head movement. In my material it was activated by verbs like *to run away*, *to jump away, to disappear*.
	words of looking/ actions involving eyes	When describing something connected to eye movements the informants sometimes illustrated them with actual eye movements Namely, with such words as *to look, to see, to imagine* (eyes closed)*, a spectator* (head movement aside and up-down gaze)*, to sleep* (eyes closed).
	eye gaze metaphors	This subtype includes some linguistic metaphors illustrated by eye gaze. 1) **Size**, even in metaphorical and implicational senses, tends to be illustrated by expanded(~big) or narrowed(~small) eyes: *significant, great, very, small, insignificant, a tiny bit.* This phenomenon can be called eye gaze metonymy as it illustrates a part of the meaning in the corresponding distance between eyelids. 2) **‘Unclear’ metaphors.** Words s*trange*, *unclear* (in ‘not understanding’ sense), and their synonyms can be illustrated by narrowed eyes, as if a person is trying to actually discern something in the space in front of them.
	imaginary object usually mapped with hands	In the contrary to usual gesticulation, when a person is mapping an imaginary object with hands when describing it - they usually mark it with gaze directed to the imaginary object mapped.
	emotions & words of emotions	This subtype refers to showing actual emotions and describing them (see 2.1.2). However, this subtype usually includes other movements than eyes
	quote	When quoting a dialogue/a statement a people tend to locate imaginary speakers in the space as if taking their

		role in a little play (resembles ‘enumeration’ function described above).
	verbs of thinking	When talking about *thinking, imagining, etc.* people sometimes look around like in the ‘thinking’ function above. This verifies conscious knowledge about ‘thinking’ eye gaze function.
accompanying gesture		‘Accompanying gesture’ means that an eye gaze is somehow included in a whole gesture. Those examples did not interest me as I cannot describe that kind of eye gaze appart their gestural context.
towards the object of conversation		When speaking about an actual object exists in the conversation space, people tend to look at that object when referring to it. However, that function appeared only in the *recart-01* recording.

To conclude, eye gaze expressions has shown to be very rich and peaky on their triggers, linguistic, psychological, cultural by their nature. It is noticeable from the Table 3 that functions are used in relatively same context. However, the fact that those eye gaze symbols get expressed and can be determined by an annotator shows that they somehow ‘compete’ between each other and some of them ‘win’, resulting in a certain eye gaze pattern.

3.5 Function - direction correspondence

However, the description of the functions is not enough to reconstruct eye gaze behavior on a machine. There is a need to detect how do those functions correspond eye gaze directions. This correspondence is shown in Tables 4 and 5.

Table 4. *Eye gaze function-direction correspondence - rec-art01 informant*

rec-art01								
Eye gaze function tag	**up**	**left**	**down**	**right**	**closed**	**object**	**eye contact**	**total number**
thinking	9%	11%	29%	51%				**133**
speaking							100%	**118**
eye gaze pattern	4%	8%	18%	34%	7%	5%	24%	**91**
listening							100%	**48**
attention to person's reaction							100%	**30**
phrasal stress							100%	**26**
softening anti-social situation			33%	67%				**18**
accompanying gesture			67%	33%				**10**
mimicking interlocutor's eye gaze						100%		**10**
joking				80%			20%	**8**
towards the object of conversation				100%				**4**

Table 5. *Eye gaze function-direction correspondence - rec-art2 informant*

rec-art02								
Eye gaze function tag	**up**	**left**	**down**	**right**	**closed**	**object**	**eye contact**	**total number**
thinking	25%	48%	14%	13%				**380**
speaking							100%	**168**
phrasal stress							100%	**114**
eye gaze pattern	15%	29%	14%	19%	11%	5%	7%	**108**
listening							100%	**50**
attention to person's reaction							100%	**38**
enumeration	13%	46%	17%	21%			4%	**24**
softening anti-social situation	5%	48%	29%	19%				**25**
joking	30%	50%	10%	10%				**10**
mimicking interlocutor's eye gaze						100%		**2**

Some of the functions have a very high percentage of a certain direction. Those include ‘speaking’, ‘phrasal stress’, ‘listening’, ‘attention to person’s reaction’ - eye contact only; ‘mimicking interlocutor’s eye gaze’ - object only. The fact that mimicking interlocutor’s eye gaze is usually towards the object simply means that it is what psychologists call *joint attention*, an important social sign of conversational engagement.

On the contrary, some of function have a more complicated directions distribution. Some of those are functions expressed by either direction: ‘thinking’, ‘enumeration’, ‘softening anti-social situation’; but some of them are complex patterns with, more or less, fixed combination of directions: ‘eye gaze patterns’, ‘joking’.

3.6 Eye gaze behavior: generalizations

Corpus analysis resulted in a little dictionary of types and subtypes of eye gaze symbols expressed in the videos analyzed (Table 3 above). Eye gaze can be extremely iconic; it reveals a lot about one’s linguistic mind map. Some linguistic metaphors can be expressed by eye gaze. Eye gaze can serve speech illustration by imitating situations, other people, situations, emotions. It can also serve social purposes showing conversational engagement, softening inappropriate behavior, flirting.

Different functions can be used in the same context, which leads to a conclusion that eye gaze system is a combination of different states. They emerge when certain internal or external stimuli trigger them in the conversational context, and they compete between each other determining which state should be expressed. The state-winner gets to be expressed by its specific set of eye gaze directions. In my system I try to reconstruct that process on the robot (Section 5).

However, the question that corpus analysis has left uncovered is whether people actually notice and pay attention to those symbols when communicating; and how important eye gaze is in human-robot interaction. I decided to answer to those questions by means of a short evaluation experiment (Section 4).

4. Eye gaze perception in the robot: preliminary experiment

The first, preliminary, experiment was based on the idea that eye contact makes human-robot conversation generally more enjoyable and human-like. The assumption was that the lack of eye contact is perceived differently in comparison to its presence. The assumption would be verified if people perceive the robot differently when listening to the same story with slightly different eye gaze behavior.

4.1 Participants

Ten participants took part in this experiment: five females, and five males. Only Russian native speakers were asked to participate, not only for the reason that both speech and surveys were presented in Russian, but because eye gaze perception and non-verbal language are proved to be culture-specific (Matsumoto, Hwang, & Frank, 2016; McCarthy, Lee, Itakura, & Muir, 2006). The mean age of the participants was 36.5 years (in the range of 24-67 years). The age scatter did not matter for the experiment as the aim was to determine robot eye gaze perception for an average person.

4.2 Design

Figure 1 shows the experiment in progress:

Figure 1. *Experimental setting*

Participants were asked to listen to a short story told by the robot twice (duration of one story - 2.5 minutes). Everything but eye gaze behavior was identical both times. While listening to the robot participant could move in the perimeter of a rectangle located on the floor (see Figure 2).

Figure 2. *The area participants could move inside*

Conditions were the following:

Condition 1 (control condition) Robot pronouncing 8 short phrases, pausing after every sentence for maximum 2 seconds long. Robot avoids eye contact with the user, looking away both when talking and in pauses between phrases (see robot looking right in Figure 3 below).

Condition 2 (experiment condition) Robot pronouncing 8 short phrases, pausing after every sentence for maximum 2 seconds long. Robot looks at the user when talking, tracking user by the means of Kinect Xbox 360 face tracker, and looks away in pauses between sentences. See robot looking at the user in Figure 4 below.

Figure 3. *Robot F-2 looking aside*

Figure 4. *Robot F-2 creating eye contact*

The order of the conditions was assigned randomly in order to avoid the influence of the novacy of the text and the robot itself on robot's evaluation.

After listening to each condition participants had to fill in an evaluation survey on their impression of the robot.

Survey. The survey consisted of 12 semantic differential scale questions about robot perception. Participants had to rank their impression of the robot on 1 to 7 scale: 1 - closer to the negative feature, 7 - closer to the positive one. The features pairs were following: cheerful - gloomy, friendly - unfriendly, pleasant - unpleasant, confident - unconfident, kind - mean, talkative - withdrawn, attentive - selfish, emotional - unemotional, calm - nervous, honest - dishonest, attractive - unattractive, fast - slow. That part of the survey aimed to define whether there is an implicit effect of eye contact in the conversation. In other words, whether unconsciously people rank eye contact condition better than control condition.

After completing both scaling parts, straight after listening to each condition, the participants were asked to answer which condition they liked better, and whether they spotted the difference between two conditions. Latter questions aimed to test the explicit difference: whether people consciously spot the presence of eye contact and how they rank conditions according to that knowledge.

4.3 Results & Discussion

Semantic differential scaling tests did not show any statistically significant results inside each parameter (Mann-Whitney U Test of 34-48; asymptotic significance (2-tailed) of > 0.3). Although the differences are not significant, there is an overall tendency to choose the condition with the eye contact over the one without it. That might indicate that the pairs of parameters were chosen not the best way. However, it also shows that implicitly people prefer eye contact strategy to the one without it.

The explicit questions showed participant's significant preference for the condition with eye gaze included (chi-square test, $p=0.04$). All the participants confirmed they could feel the difference in the robot's body language. However, only one participant identified that the difference was in the eye contact.

In conclusion, the preliminary experiment confirmed that people do notice differences in robot's eye gaze behavior. They do not realise what that difference and they tend to interpret it differently, but they still give their preference to the robot that maintains eye contact with them. However, those eye gaze strategies were constructed manually and would not fit all the contexts. Thus, the next step was to construct meaningful eye gaze moderating system based on the corpus analysis.

5. Modelling eye gaze in robot

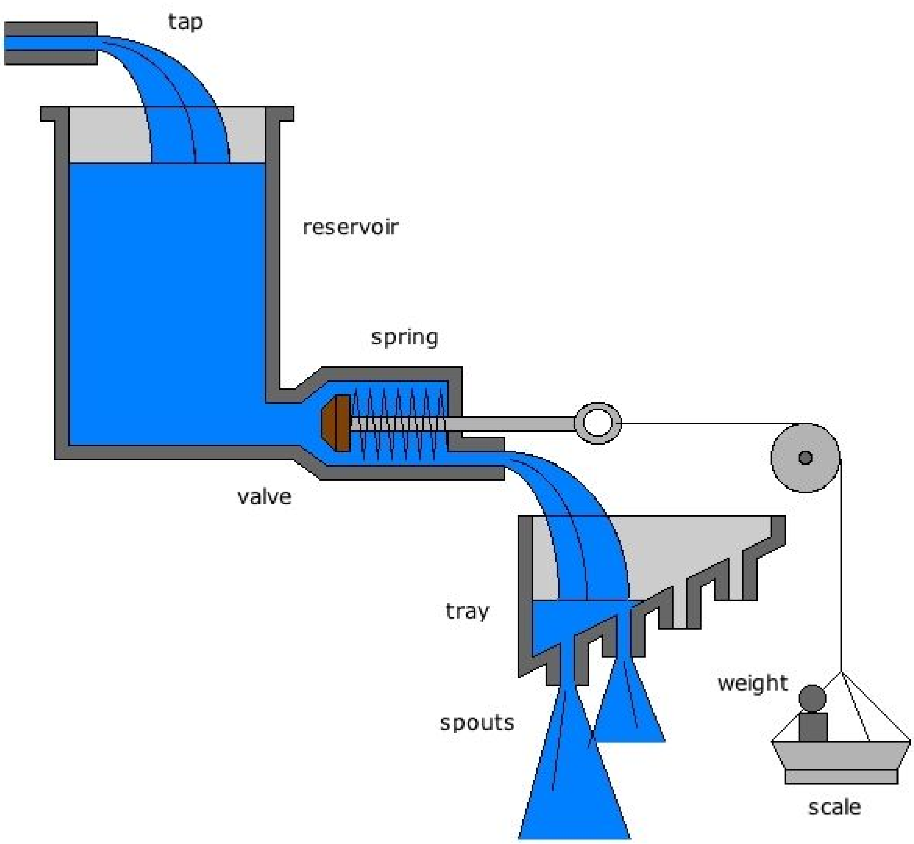
After determining a set of oculomotor movements and their distribution, and verifying their significance for humans, my aim was to design them in the robot. The technologies I used to model robot's eye gaze behavior included Python, Behavior Markup Language (BML), and an external face tracker Kinect Xbox 360, recognising user's eyes and face location.

5.1 Theoretical foundation: assumptions

As a result of corpus data processing, I made certain assumptions my program is based on. My hypothesis was that there are certain states evolving in human brain when communicating. Activation of those states depends on internal and external stimuli. Every state has its way of activation change.

This assumption is partially based on K. Lonrenz's hydraulic behavioral model (Lorenz 1963). Although his theory is primarily about instinctive animals' behavior, it perfectly suits humans at the same time. In Lorenz's words, behavior resembles hydraulic system shown in Figure 5.

Figure 5. *K. Lorenz's hydraulic model (Lorenz 1963)*



Internal motivation produce water from the tap, which fills up a reservoir. The water from reservoir pushes a spring out in a certain extent. At the same time there are external stimuli symbolically represented by weights on a scale, which pull the spring as well. The spring releases water to a tray with holes - the more water is released at the same time, the more 'spouts' it fills. The number of spouts represents the strength of reaction expressed. It can be illustrated by a goose hunger: the growing hunger is an internal factor filling up a reservoir. When a goose notices a piece of bread it adds another external factor on the scale. When spring is pushed away, the water will be released and the goose will eat the bread in an extent of passion depending on how hungry it was or how attractive the bread is.

Similarly, in my program, different external and internal factors push the spring out and release the water from reservoir increasing state's activation. At each period of time states compete in their level of activation. The winner state sends a certain eye gaze to the robot's interface. The more 'water flows out' the higher is the probability of a state to send its gaze to the machine.

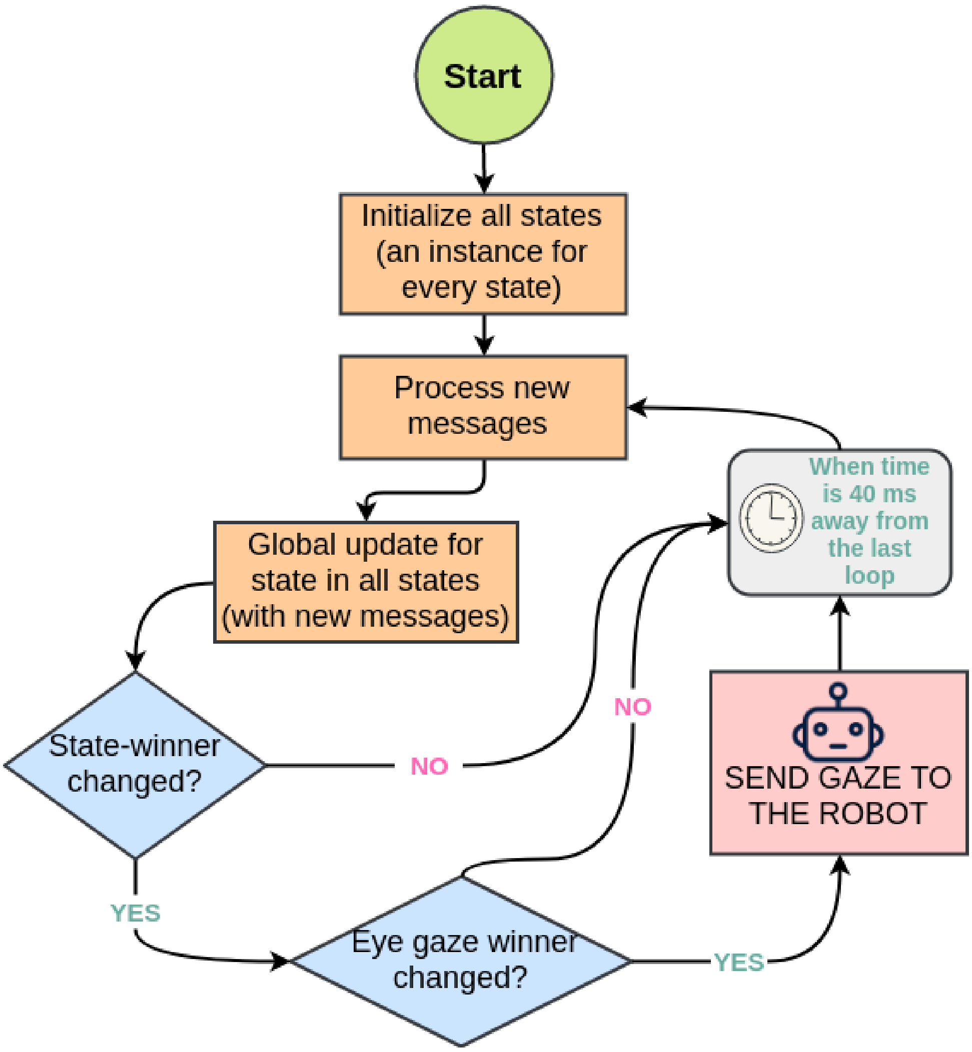
The states competing in my system are based on the functions described above, in Section 3.2.2. Every state has its own mathematical function which changes its activation value when receiving certain messages from the robot's system. Those messages, representing those internal and external factors, affect states according to the algorithms assigned.

5.2 Program structure

My model is built in Python and is currently availible in a github repository here: https://github.com/mmtsfasman/Eye_Gaze_System_for_F-2_Robot.

The program itself is started through `core.py` script and works by an algorithm symbolically represented on the chart 6 below.

Figure 6. *Program functioning algorithm (simplified)*



The program listens to certain messages received from the robot's system through a network socket connection. My script processes those messages and changes state's activation correspondingly, if no messages are received my program still updates the states, but without any new changes in states' mathematical functions. The state with the greatest activation 'wins' and if it is not the same state as the previous winner and it is not the same eye gaze direction, it sends a bml package with its gaze to the robot. It all happens in 40 milliseconds loop similar to 25 frames per second in cinema or cartoon animation.

Bml packages are what robot's behavior is built on, they are strings with a certain syntax, and are structured as in the following example of eye gaze in up-left direction:

```
<bml id="22" syncmode="single">
      <head id="2" lexeme="eyes_up_left3"/>
      <pupils id="3" lexeme="eyes_up_left3"/>
</bml>.
```

Script `utils.py` stores secondary functions used in the core and states scripts.

5.2.1 Message receiving and processing

So far the incoming messages from the robot's system include the following:

- ➔ `'Turn completed.'` - when robot has completed bml package to look aside.
- ➔ `'Gaze completed.'` - when robot has completed bml package to look towards a person (creating eye contact)
- ➔ `'"`*`text`*`" started at `*`time`*`. Stroke delay `*`time`*`.'` - when a bml package with speech has started playing on the robot. *Text* is a string with robot's speech, start *time* is the time robot starts pronouncing the speech, delay *time* is the time period after which phrasal stress is expected in milliseconds.
- ➔ `'Speech completed.'` - when speech bml package is finished.

Messages are processed into a set of stimuli in the right format for the states to use through `process` class stored in `proccess.py` script in the repository.

5.2.2 States classes hierarchy

States based on the function system described in Section 3.2.2 are realised as a system of Python classes[1]. The parent class is called `state` and includes all the functions and attributes all states have. Individual states are `state`'s daughter classes. Eye gaze function classification from 3.2.2 Section is slightly reorganized so that it could be implemented most effectively. Function - implementation correspondence is shown in Table 6. Some of the functions are not implemented due to the limits of robot's recognition, and speech processing

[1] This part of the script is stored in stored in `states.py` file availible here: https://github.com/mmtsfasman/Eye_Gaze_System_for_F-2_Robot/states.py

abilities. They can be implemented later, when certain information about the interlocutor position and behavior, speech characteristics, robot emotions is automatically extracted.

The functions implemented, partially implemented and not yet implemented are shown in Table 6:

Table 6. *Eye gaze function's implementation*

Function	Implemented?	How implemented? (Or why not implemented?)	What class?	Subtypes inside that class
thinking	yes, fully	$f = Cos(x)$ Independent from any internal/external factors or incoming messages function. Created so due to the fact that no human processes can be represented as a constant line - it usually can be approximated to a pulsing curve.	`thinking`	*no extra factors*
speaking	yes, fully	$f = -ax^2 + bx + c$ Where *a, b, c* are predicted every time a speech startpoint and phrasal stress point are determined. Parabola, with the peak in the phrasal stress, and abscissa intersection points in the starting and ending point of robot's phrase.	`speaking`	✔phrasal stress: determines the peak of the parabola
attention to person's reaction	yes, fully	$f = (x - a)/c$ Where a is time in a moment when robot looked towards the user, so that it grows from zero every time robot looked user in the eye.	`attention_to_person`	✔question: after robot asks a question the activation jumps on x
softening anti-social situation	partially	Still requires automatic speech processing.	`anti_social`	✔disagreeing ✘ too long eye contact ✘ softening 'strong' words
remembering	partially	Still requires automatic speech processing do that factors are determined.	`remembering`	✔argumentation ✘ remembering a word/a situation ✘ mental arithmetic

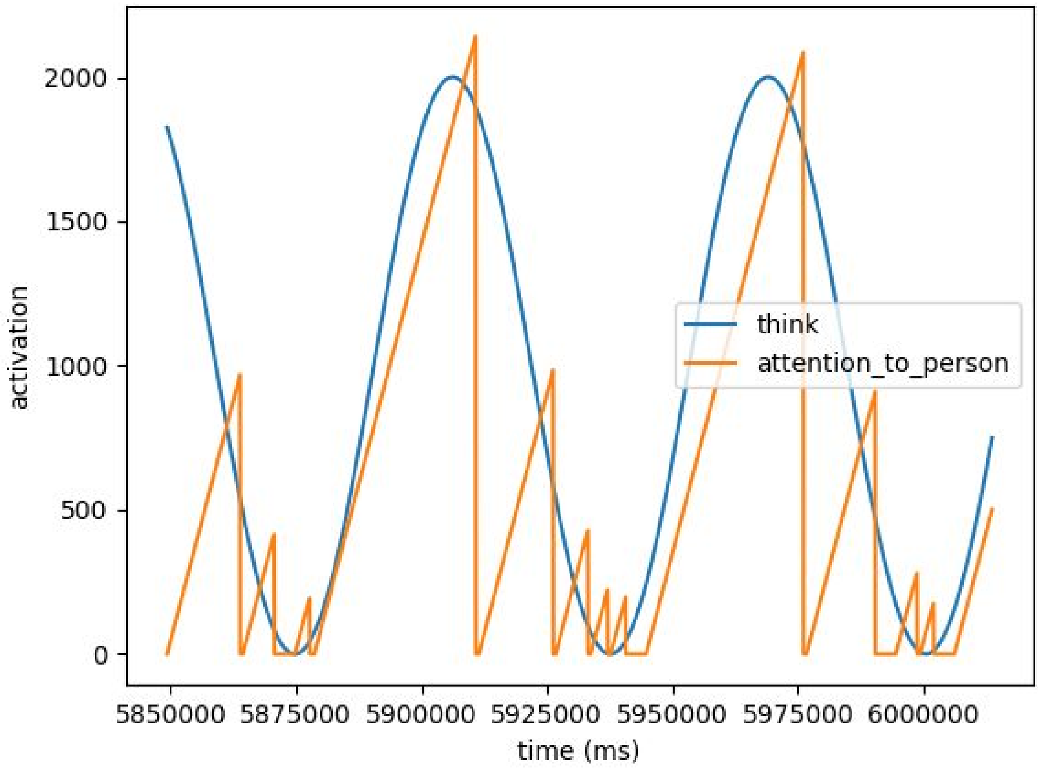
				✘ thinking on the answer
agreeing	partially	Requires eyelids in robot's interface (to close eyes). However, if the bml package with closed eyes is added - my class would function. It will be x^2 function.	`agreeing`	*no extra factors*
joking	*no*	Requires automatic joke recognition.		*no extra factors*
listening	*no*	Requires automatic sound source recognition (to determine who is speaking).		*no extra factors*
mimicking interlocutor's eye gaze	*no*	Requires eye gaze direction recognition (where is interlocutor looking?).		*no extra factors*
eye gaze patterns (iconic eye gaze)	*no*	Requires special speech processing and determination of triggering words.		✘ quote class ✘ verbs of movement & thinking ✘ words of looking ✘ eye gaze metaphors ✘ imaginary object usually mapped with hands ✘ emotions & words of emotions

5.3 Examples of system functioning on the robot

When the system is interrupted, it plots the activation functions for every state that has been activated during the system functioning period.

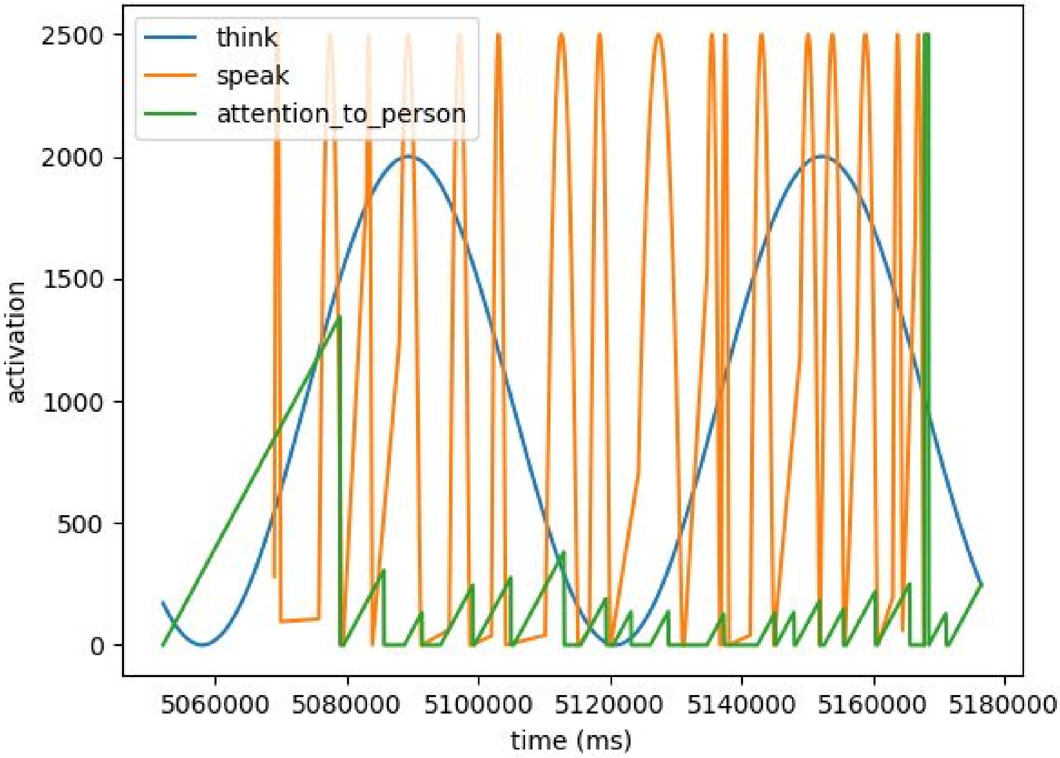
The first example, Figure 7, shows how the system is functioning by itself, without any extra robot's actions. In other words, if the system is on and robot is in the state of inaction, robot's eye gaze behavior changes by the functions shown on the picture. Think state is a cosine function (blue curve) independent from any stimuli. Attention to person is a linear function (orange lines) function that changes when the eye contact is achieved: it falls to zero and begins its new linear function from that point. The points where a new state beats all the other ones are the points of that state's victory - the points of sending eye gaze bml package to the robot.

Figure 7. *Example of states activation in robot's inaction*



As opposed to the state of inaction (Figure 7 above), the second example (Figure 8 below) represents robot's eye gaze behavior when speaking.

Figure 8. *Example of states activation when robot is speaking*



Attention to person state's sudden jump in the end of the phrase indicates that the phrase was a question and the robot came back to eye contact in order to listen to the answer or to check out interlocutor's reaction triggered by that question. Parabolas (orange curves) evolve when robot is speaking.

Figure 9. *Example from figure 8 with phrase limiting points illustration*

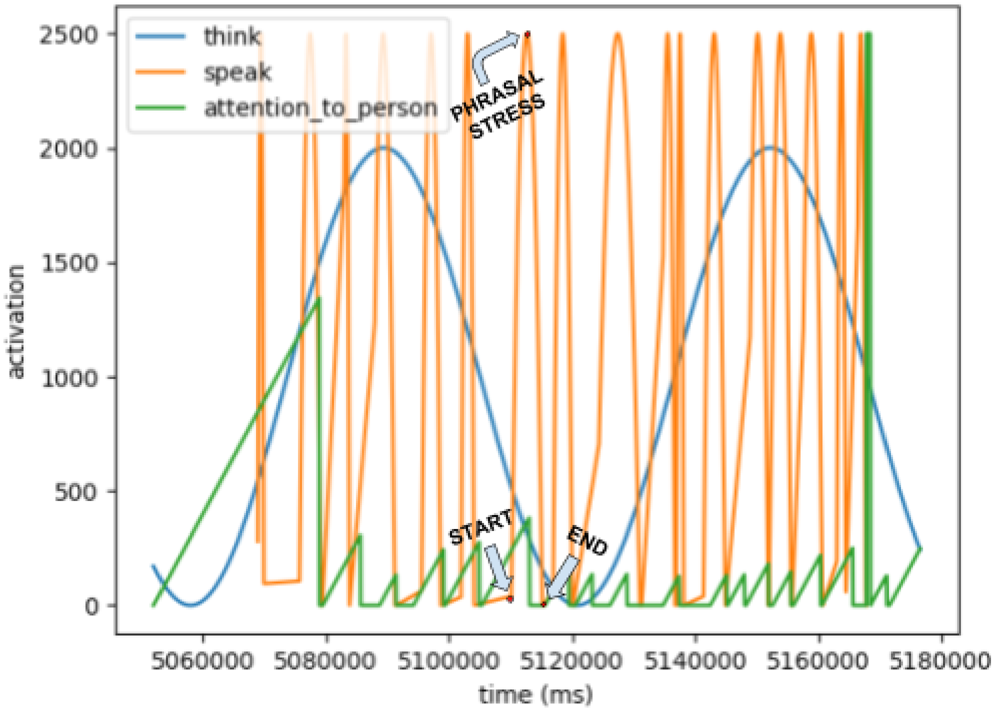


Figure 9 illustrates that phrases starting and ending points can be spotted by the points parabolas intersect horizontal axis (see start point and end point in Figure 9). Parabola's peaks, on the other hand, are phrasal stress moments (see phrasal stress point in Figure 9).

5.3 Generalizations

The suggested model is constructed on three principles. First, it appeals to the observations of human eye gaze behavior based on the Russian Emotional Corpus. Second, on the hypothesis that internal states compete between each other, and the winner state gets expressed. Third, on K. Lorenz's hydraulic model. In order to verify whether this model reached the aim to increase robot's likeability and believability, we tested it in human-robot interaction experiment (Section 6).

6. Main experiment: user experience testing

The concluding experiment aimed to test the designed model and users' perception of the robot when the system is on. The main question the experiment was to answer: how does the model affect user perception of the robot, its likeability and believability? The structure and main concept of this experiment was the same as the preliminary one described above in Section 4. However, the conditions were different, and the survey was improved and expanded.

6.1 Participants

14 Russian native speakers took part in the experiment: six females, eight males. The mean age of the participants was 26.8 years (in the range of 21-45 years).

6.2 Design

Figure 10. *Experimental setting*

Similarly to the preliminary experiment, participants were asked to listen to the robot telling one story twice (duration of one story - 2.5 minutes). The participant standing in front of the robot throughout the experiment so that they notice the contrast between a regular sight and a sight towards them. They were allowed to move around a rectangular mapped out with a colorful tape on the floor.

The only way two conditions differed from one another was that in the **test condition** my model controlled robot's eye gaze, in the **control condition** the robot looked to the right or to the left seriatim every 6 seconds. In the test condition, my script was turned on. It was fully functioning: the states were evolving when certain messages from robot's system were receives; competing between each other; the winner state would send a package right to the robot's behavioral queue with the highest weight, which means it was always the one to get expressed.

In order to avoid the effect of novacy of the story half participants listened to stories in test condition - control condition order, and the other half - the other way around.

Survey was meant to test the difference between users' perception of the robot with my system on and the side-to-side gaze. In the **first** part, participants had to rank their impression of the robot using semantic differential scales. It was meant to test implicit difference in their perception of the conditions. Feature pairs tested through the evaluation form can be divided in three groups:

1) movement quality: calm - fussy, fast - slow
2) robot's attitude to the user: friendly - aggressive, attentive - inattentive, attractive (pleasant) - unattractive (unpleasant)
3) impression of the robot: reflective (word *zadumchivy* in Russian) - spontaneous, emotional - emotionless

Participants were asked to fill out that part twice straight after listening to each condition. After listening to both condition and evaluating those qualities, participants had to complete the **second**, explicit comparison, part of the form. It included 7 questions about the same qualities as in the semantic differential part, but in comparison form. The questions were in 'Which time robot seemed (more calm/faster/friendlier/more attentive/more attractive/more reflective/more emotional) to you?' form. There were three options for those questions - 'the first', 'the second', 'didn't seem to differ'.

6.3 Results

Semantic differential scales did not show any significant dependencies. However, same as in the first experiment, there was a general trend of giving higher ranks to the robot with the eye gaze generating system on.

Explicit comparison part clearly showed users' preference for the robot with the eye gaze system constructed throughout this project: in the test condition the rates appeared to be significantly better than in the control one. Table 7 indicates the results for each parameter, the insignificant parameters are highlighted grey:

Table 7. *Main experiment's results*

parameter	**result**	**chi-square test, p-value**
attentiveness	12 out of 14 (85%) participants considered the robot with my model more attentive than the other	highly significant ≤ 0.01 (0.0008)
friendliness & emotionality	11 out of 14 (78%) participants considered robot controlled by my system more friendly and emotional than the random one; remaining three participant did not spot the difference in this aspect	highly significant ≤ 0.01 (0.003)
attractiveness	9 out of 14 (64%) participants considered robot with my model more attractive than the other, remaining 5 did not spot the difference in this aspect	highly significant ≤ 0.03 (0.029)
calmness and speed	parameters appeared to give random rates - no significant correlation was spotted at this point	not significant > 0.2 (0.29 & 0.22 correspondingly)
reflectivity	6 participants considered robot in the control condition more reflective, 5 participants did not spot the difference, and only 3 considered robot with the designed system more reflective. Such distribution may be due to the fact that people tend to search for explanation for any behavior, even if it is completely random. In this case people interpreted right-left gaze as a robot in a 'thinking mode'.	not significant > 0.2 (0.22)
general likability	The answers to 'Which robot behavior did you like better?' question appeared to show generally more preferences for the robot with my model: only 2 of 14 participants liked the control condition better than my model.	highly significant ≤ 0.01 (0.004)

6.4 Discussion

The experiment verified the hypothesis that robot with eye gaze controlled by the constructed system is perceived as generally more friendly, attentive, emotional, and attractive. First, this confirms to that the aim of making F-2 robot more likable and believable is achieved. Second, it confirms that meaningful eye gaze system is perceived as more communicationally appropriate even in human-robot interaction. The designed model can be used to control eye gaze behavior of anthropomorphic robots, communicating with humans.

7. Conclusions

My project implements a two-way interaction between theory and the practice of AI development: the latter proceeds from a theoretic model into an AI product which, in its turn, can suggest corroborative evidence to the model. On the one hand, I use corpus analysis in order to construct modules for a conversational agent system. On the other hand, I use a conversational agent system to evaluate how the resulting model of eye expressions affect robot's perception in human-agent interaction.

To my knowledge, the resulting system, is the first to view eye movement design in a robot from the perspective of communicational significance of certain eye expressions, as opposed to systems that primarily focus on anatomical resemblance and oculomotor apparatus physiology.

As a result of my work I presented a description of meaningful eye gaze behavior in a dialogue. The eye gaze proved to serve social, linguistic, and illustrational purposes. My 'dictionary' included 14 groups and 16 subgroups of communication eye gaze symbols. The list can still be expanded if processing more video material of different conversational situations. The results may be specific for Russian cultural and linguistic behavior, and it would be particularly interesting to analyse eye gaze behavior of people from different cultures in the further research; and test how people of different cultures perceive the robot with my system moderating its eye gaze behavior.

The suggested computer model is implemented in an F-2 Emotional Agent, improving its communicational and empathy-triggering skills. It is based on the idea of states competing between each other, and sending an eye gaze to the robot when winning. Not all the states discovered by the means of corpus analysis are implemented in F-2 Emotional Robot system yet. However, as user experience testing has shown, users already perceive the robot as more attractive, attentive, friendly, and emotional when the resulting model moderates robot's eye gaze.

Although the code is specific for that model of the robot, the outcome of my work can be used as a theoretical basis to design eye gaze communication on various conversational interfaces. Moreover, my results are of relevance to the study of human non-verbal communication specifics, as I suggest an alternative classification of conversational eye gaze behavior.

8. References

Admoni, H., & Scassellati, B. (2017). Social Eye Gaze in Human-Robot Interaction: A Review. Journal of Human-Robot Interaction.

Admoni, H., Hayes, B., Feil-Seifer, D., Ullman, D., & Scassellati, B. (2013). Are you looking at me? Perception of robot attention is mediated by gaze type and group size. In ACM/IEEE International Conference on Human-Robot Interaction.

Al Moubayed S., Edlund J., Beskow J. (2012) Taming Mona Lisa: Communicating gaze faithfully in 2D and 3D facial projections, ACM Transactions on Interactive Intelligent Systems (TiiS), v.1 n.2, p.1-25

Altmann, G. T. M. (2012). The mediation of eye movements by spoken language. In The Oxford Handbook of Eye Movements.

Andrist S., Pejsa T., Mutlu B., Gleicher M. (2012) A head-eye coordination model for animating gaze shifts of virtual characters, Proceedings of the 4th Workshop on Eye Gaze in Intelligent Human Machine Interaction, p.1-6

Bee, N., & André, E. (2008). Cultural gaze behavior to improve the appearance of virtual agents. International Journal of Psychology.

Bee, N., André, E., & Tober, S. (2009). Breaking the ice in human-agent communication: Eye-gaze based initiation of contact with an embodied conversational agent. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics).

Bee, N., Wagner, J., André, E., Vogt, T., Charles, F., Pizzi, D., & Cavazza, M. (2010). Discovering eye gaze behavior during human-agent conversation in an interactive storytelling application. In International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction on - ICMI-MLMI ’10.

Bérard P., Bradley D., Nitti M., Beeler Th., Gross B. (2014) High-quality capture of eyes, ACM Transactions on Graphics (TOG), v.33 n.6, November 2014

Blaison, C., Hareli, S., Strauss, C., & Hess, U. (2012). Imitating eyes: Mimicry of emotions shown in the eyes. Paper presented at CERE 2012, Canterbury, England.

Burgoon, J. K., Coker, D. A., & Coker, R. A. (1986). Communicative Effects of Gaze Behavior. Human Communication Research.

Campana, E., Baldridge, J., Dowding, J., Hockey, B. A., Remington, R. W., & Stone, L. S. (2001). Using eye movements to determine referents in a spoken dialogue system. In Proceedings of the 2001 workshop on Perceptive user interfaces - PUI ’01.

Colburn, R. A., Cohen, M. F., & Drucker, S. M. (2000). The Role of Eye Gaze in Avatar Mediated Conversational Interfaces. Microsoft Research Report.

Das, A., & Hasan, M. M. (2014). Eye gaze behavior of virtual agent in gaming environment by using artificial intelligence. In 2013 International Conference on Electrical Information and Communication Technology (EICT). https://doi.org/10.1109/EICT.2014.6777879

Francois G., Gautron P., Breton G., Bouatouch K. (2009) Image-Based Modeling of the Human Eye, IEEE Transactions on Visualization and Computer Graphics, v.15 n.5, p.815-827

Grishina, E. A. (2005) Dva novyh proekta dlja Nacional'nogo korpusa: mul'timedijnyj podkorpus i podkorpus nazvanij // Nacional'nyj korpus russkogo jazyka: 2003—2005. M.: Indrik, 233—250.

Grishina E. A. (2012) Ukazanija rukoj kak sistema (na materiale Mul'timedijnogo russkogo korpusa) // Voprosy jazykoznanija. No 3. (Гришина Е. А. Указания рукой как система (на материале Мультимедийного русского корпуса) // Вопросы языкознания. 2012. No 3.)

Gu E., Lee S., Badler J. B., Badler N. I. (2007) Eye movements, saccades, and multiparty conversations. In Data-Driven 3D Facial Animation. Springer-Verlag London Limited, pp. pp.79-97.

Hacker, C., Batliner, A., & Nöth, E. (2006). Are you looking at me, are you talking with me: Multimodal classification of the focus of attention. Proc. of the 9th International Conference on Text, Speech and Dialogue (TSD 2006).

Ishii, R., & Yukiko, I. N. (2010). An Empirical Study of Eye-gaze Behaviors : Towards the Estimation of Conversational Engagement in Human-Agent Communication. EGIHMI ’10 Proceedings of the 2010 Workshop on Eye Gaze in Intelligent Human Machine Interaction.

Ishii, R., Nakano, Y., & Nishida, T. (2013). Gaze awareness in conversational agents: Estimating a user’s conversational engagement from eye gaze. ACM Transactions on Interactive Intelligent Systems. http://doi.org/10.1145/2499474.2499480

Kibrik A., Budennaya E., Buryakov M., Dobrov G., Evdokimova A., Fedorova O., Korotaev A., Litvinenko A., Nikolaeva Yu., Podlesskaya V., Sukhova N., Zherdev I. (2018) Russian Multichannel Discourse - Institute of Linguistics, Russian Academy of Sciences. URL: http://www.multidiscourse.ru

Komogortsev O., Holland C., Jayarathna S., Karpov A. (2013) 2D Linear oculomotor plant mathematical model: Verification and biometric applications, ACM Transactions on Applied Perception (TAP), v.10 n.4, p.1-18

Kotov, A., Budyanskaya, E. (2012) The Russian emotional corpus: communication in natural emotional situations. In: Computational System and Intellectual Technologies, vol. 11(18), pp. 296 - 306. RSUH, Moscow

Kreysa, H., & Pickering, M. J. (2012). Eye movements in dialogue. In The Oxford Handbook of Eye Movements.

LaFrance, M., & Vial, A. C. (2016). Gender and nonverbal behavior. APA Handbook of Nonverbal Communication.

Loomis, J. M., Kelly, J. W., Pusch, M., Bailenson, J. N., & Beall, A. C. (2008). Psychophysics of perceiving eye-gaze and head direction with peripheral vision: Implications for the dynamics of eye-gaze behavior. Perception. http://doi.org/10.1068/p5896

Lorenz, K. (1963). On aggression. On aggression. Harcourt, Brace and World: New York.

Luck, S. J. (1998). Neurophysiology of Selective Attention. Attention.

Macrae, C. N., Hood, B. M., Milne, A. B., Rowe, A. C., & Mason, M. F. (2002). Are You Looking at Me? Eye Gaze and Person Perception. Psychological Science.

Matsumoto, D., Hwang, H. C., & Frank, M. G. (2016). APA handbook of nonverbal communication. APA handbooks in psychology.

Mori, M. (1970). The uncanny valley. Energy, 7(4)

Pejsa, T., Andrist, S., Gleicher, M., & Mutlu, B. (2015). Gaze and Attention Management for Embodied Conversational Agents. ACM Transactions on Interactive Intelligent Systems. https://doi.org/10.1145/2724731

Pelachaud, C., & Bilvi, M. (2003). Modelling Gaze Behavior for Conversational Agents. Intelligent Virtual Agents.

Scherer, K. R. (1988). On the symbolic functions of vocal affect expression. Journal of Language and Social Psychology, 7(2), 79–100.

Sibert, L. E., & Jacob, R. J. K. (2000). Evaluation of eye gaze interaction. In Proceedings of the SIGCHI conference on Human factors in computing systems - CHI ’00.

Tang, D., & Schmeichel, B. J. (2015). Look Me in the Eye: Manipulated Eye Gaze Affects Dominance Mindsets. Journal of Nonverbal Behavior. http://doi.org/10.1007/s10919-015-0206-8

Torres, O., Cassell, J., & Prevost, S. (1997). Modeling Gaze Behavior as a Function of Discourse Structure. First International Workshop on Human-Computer Conversations. Bellagio, Italy.

Yagi, T. (2010). Eye-gaze interfaces using electro-oculography (EOG). In Proceedings of the 2010 workshop on Eye gaze in intelligent human machine interaction - EGIHMI ’10.