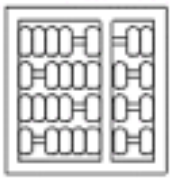

INSTITUTO DE COMPUTAÇÃO
UNIVERSIDADE ESTADUAL DE CAMPINAS

Virtual Presenter and *Vila na Rede*: augmenting accessibility in ISNs

Elaine C. S. Hayashi *Leonelo D. A. Almeida*
M. Cecília C. Baranauskas

Technical Report - IC-10-019 - Relatório Técnico

May - 2010 - Maio

Virtual Presenter and *Vila na Rede*: augmenting accessibility in ISNs

Elaine C. S. Hayashi, Leonelo D. A. Almeida, M. Cecília C. Baranauskas

{hayashi, leonelo.almeida, cecilia}@ic.unicamp.br

Abstract. *Vila na Rede* is an Inclusive Social Network (ISN) system developed as a joint effort with Brazilian communities. As a product of *e-Cidadania's* Project, it has inherited the objective of being accessible for the widest variety of users, including those less familiar with technology or with low literacy levels. In this direction, different features were incorporated into *Vila na Rede* in order to provide its users with scaffolding resources that help them profit more from the system. One of these features is the Virtual Presenter, a talking head that allows users to have the textual information converted into speech, presented by a face that moves its lips accordingly. This technical report provides details on the integration of the Virtual Presenter into *Vila na Rede*, as well as the activities that were conducted to evaluate this new feature at the ISN.

Keywords: Inclusive Social Networks, talking head

1. Introduction

E-Cidadania is a Brazilian research project that has taken the challenge of developing systems that allow the access and that make sense to the community of users, contributing to the constitution of a digital culture and respecting the diversity of the population. As a result, the project has launched *Vila na Rede*, an Inclusive Social Network (ISN) system, designed with, by and for Brazilian people. *Vila na Rede* is being used by citizens all over Brazil and has even been accessed from abroad, announcing ideas, goods and other initiatives from different communities.

The design of systems that make sense and that are accessible to citizens require a rather socio-technical view of the problem. Because of that, the research has adopted as a methodological reference, the principles and concepts of both Participatory Design (Mumford, 1964) and Organizational Semiotics (Stamper, 1988). *E-Cidadania's* researchers have been successfully involving end users and developers in Semio-Participatory Workshops (Neris *et al.*, 2009, Hayashi and Baranauskas, 2009) in order to construct the ISN.

In previous encounters with our target users, we have understood how the figure of an avatar in the system would be beneficial (Hayashi & Baranauskas, 2008). That workshop had the objective of analyzing how users made sense of different multimedia outputs to retrieve information. Users were grouped to perform a task in a role play situation in which they had to find the answer for a question. Each group was exposed to a different way of searching the answer. One group had access to the information in written format. Another

group was able to listen to the information, as if they were facing an Automated Response Unit (ARU), like those found in call centers. The third group had the same information available in images. The last group consulted the information from a real person, like in an information desk. The group with best results was the last group. During the discussion with all groups, they reported their wish to have similar attendants in the systems - that is, the figure of a person to support them in their online tasks.

In another activity, we could realize that computer synthesized voices would probably be well accepted by our target users. In that experiment - which was in fact related to another project and scope - users thought that the recording of the voice of a foreign researcher (speaking in Portuguese - the users' mother language, but with accent) was a synthesized voice. They enjoyed the voice and had fun with it, and most important, they were able to understand it. During yet another encounter, users accidentally had access to a feature that was being built at *Vila na Rede's* test environment and that was not ready for use. This feature - as it was implemented at that time - was able to synthesize text extracts into speech. Even though the final objective of this feature was different, since users were excited about the possibility of having the text read for them, we decided to implement the feature as it was.

Such new feature can be referred in the literature as a “talking head". In the next section, we present a brief review on such systems. In Section 3, we describe the process of implementing the existing system at *Vila na Rede*. Section 4 narrates the Semio-participatory workshop in which the feature was tested and Section 5 discusses the results from this experiment. Section 6 concludes this technical report.

2. Talking head systems

Talking heads are computer-generated facial animations systems that, by synchronizing synthesized speech with the movement of the lips of the facial animation, are capable of reproducing face-to-face communication. There are many ongoing researches on this field. For example, the iFace (DiPaola and Arya, 2007), which was designed for games and interactive learning applications. IFace's architecture models the face's personality along with its different moods. This means that the face should be able to frown, blink, and make other facial movements, in contrast with the Virtual Presenter, which can move its lips. We believe that moving the mouth would be enough for the purpose of amplifying the accessibility. Since it was thought for games, iFace's end users are already expected to have hardware that is capable of performing the computing required to support the system. Although emotionally expressive talking heads would certainly enrich the ISN even more, our target users would hardly count on enough hardware and Internet bandwidth to run such robust applications.

Other examples of talking heads can be found in the literature, like the approach proposed by Cao *et al.* (2005), which can generate matching and expressive facial movements from a given sound input; *Mike Talk* (Ezzat and Poggio, 1998), which uses Festival for converting text into speech; SynFace (Salvi *et al.*, 2009), which is available in four different languages - except Portuguese; and, which seems to be the first initiative in the subject, dating back to 1972: (Parke).

Going far beyond the scope of our work, as an example of even more sophisticated interactive faces, we can mention the virtual receptionist from Bohus and Horvitz (2005). It allows interactions in the open world - that is, outside the limits of the internet or computer systems. This receptionist is able to recognize faces in order to direct its face to the interlocutor and it is also aware of the environment (*e.g.* if a new user is approaching). The system's artificial intelligence let users talk directly to it as if it was a real person.

More realistic to our context, another example is the Virtual Presenter (Costa, 2009). Virtual Presenter is a 2D facial animation system that is integrated with a text-to-speech synthesizer designed specifically for Brazilian Portuguese. Being developed for the same language and for similar contexts of use, this talking head seemed a good alternative for the use at *Vila na Rede*. Its architecture - which is further described in the next section - also contributed to the choice.

Virtual Presenter Core (VPC) is a product of Costa's mastering thesis (Costa, 2009) developed at the School of Electrical and Computer Engineering at University of Campinas, Brazil. VPC is the internal name we used in this research to designate the Costa's prototype for 2D facial animation. Costa's research consisted of the proposition of a method for realistic reproduction of visible speech articulatory movements, including co-articulation effects, and the possibility to implement the method also on limited processing and memory platforms, like mobile phones or personal digital assistants. The developed method is based on an image database of Brazilian Portuguese context dependent visemes and uses the morphing between visemes strategy as facial animation synthesis technique. In this research we integrated the prototype of Costa's research to the ISN *Vila na Rede*.

3. The Virtual Presenter at *Vila na Rede*

The architecture built to support the integration of VPC in our ISN is based on making virtual presentations videos as services for the ISN *Vila na Rede*. Figure 1 illustrates the architecture using two servers, the first hosting the ISN and the other performing the processing of VPC and hosting the resulting videos. When new or updated data is posted from the user interface to the ISN the text content is submitted to a text-to-speech server - in this case Festival (University of Edinburgh, 2010). The text-to-speech server generates a temporized phoneme file and audio file. After the synthesis process, the ISN makes a request to the Virtual Presenter Service (VPS). So VPS starts by getting the phoneme and audio files from the ISN server and calling the VPC. The VPC prototype is a C++ application that uses Open Computer Visual library (2010). It uses a database of only 34 images and the temporized phoneme file to generate the visemes. Currently the VPC generate animations considering approximately 30 frames per second. Each frame is an image of 320 x 240 pixels. At end of the VPC execution, VPS calls FFmpeg (2010) audio and video library to join and synchronize the visemes and the audio files and storage the resulting video. Finally, VPS informs the ISN that the virtual presentation is available for the posted content.

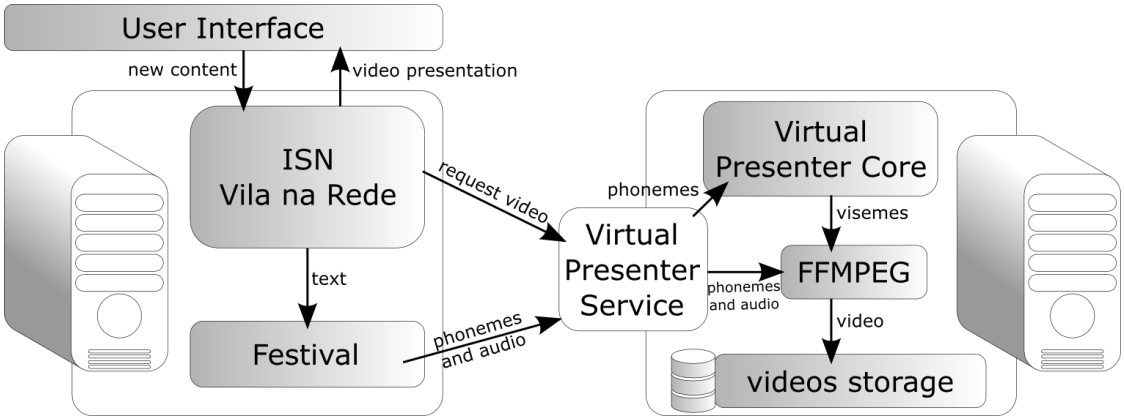

Figure 1. Virtual Presenter architecture integrated to the ISN Vila na Rede.

In the user interface layer the process for generating video presentations is transparent. When a person posts some content it is not necessary to explicitly request the creation of the video nor wait until it is generated. After the system finishes the video generation, an icon indicating the video availability is displayed along with the content, as presented in Figure 2.a. By clicking on the icon the ISN loads the video and presents it at the right region of the website (see Figure 2.b). To present the video the ISN uses the JWPlayer (Long Tail ad Solutions, 2010).

Figure 2. Virtual Presenter in the ISN *Vila na Rede*.

By the time this research was developed there was not a Festival extension for Brazilian Portuguese language that offered a female voice. To overcome this issue we changed the pitch of the male voice to obtain a neutral one. This modification did not affect the length of the video nor the phonemes pronunciation but made the voice more artificial than the original configuration.

4. A Semio-Participatory evaluation of the Virtual Presenter in Vila

In this section we describe the activities that took place in order to evaluate the new functionality regarding the Virtual Presenter at *Vila na Rede*. The three main aspects to be evaluated were: 1. Festival, considering the quality of the voice, its timbre and speed of the

speech; 2. Virtual presenter and the video x speech synchronism, how close to reality the videos are, and the role of the video in helping to understand the audio; 3. How both were implemented at *Vila na Rede*, considering the size of the video, speed and the work done to incorporate it into the website.

For this activity, we used a computer laboratory at the University of the researchers affiliations. Each participant from the community of users was assigned either to one computer or one laptop, and they all received head phones to listen to the audio and they were all able to see the videos. Each two participants were assigned to one researcher who had the role of observing their behavior. The observers were in charge of playing the announcements in the order that was specified.

The announcements that were shown to participants were previously chosen by researchers. In total, participants saw and/or listened to four announcements. Two were shown with video and sound and other two announcements had only the audio. This control was made in order to provide us with means to analyze the results. By comparing the level of comprehension we would be able to deduce the influence of the media.

From the four announcements presented to the users, two were rather short ones, intended to get a word by word transcript. The other two were longer announcements and participants were supposed to simply inform the overall idea, and not the literal transcription. Table 1 shows the announcements that were chosen and shown to the participants, as well as announcement with only the audio and with video and audio.

Table 1. Announcements used in the activity and its form and content.

Link	Media	Content (in Portuguese)
Shorter announcements:		
http://www.vilanarede.org.br/node/855	audio	bordados fita, ponto cruz e outros. Preços combinar.
http://www.vilanarede.org.br/node/974	audio + video	boneca de fuxico. Essa é uma bonéca feita de fuxico.
Longer announcements:		
http://www.vilanarede.org.br/node/1191	audio	Nova funcionalidade no Vila: a Apresentadora Virtual. Em breve teremos a Apresentadora Virtual que lerá os anúncios para você. O lançamento será em março.
http://www.vilanarede.org.br/node/938	audio + video	Festa Junina SAMUCA. SAMUCA - Rua Antonio Provatti 301 - Jd .Triunfo. Venham Prestigiar e Participar da Grande Festança do Arraia do SAMUCA, teremos

		barracas típicas e muito brindes para todos, não percam.

Right after seeing the announcement, participants completed a form with the information asked. As mentioned before, for the first two announcements, participants were supposed to write in this form the actual content heard; and for the last two, the general idea understood.

4.1 Results

The results obtained from the forms were analyzed considering two metrics. First, considering the two shorter announcements, we counted the substantives and verbs from them *i.e.*, for the node 855, there are 7 words and, for the node 974 there are 5. Second, as the participants were supposed to write what they had understood about the longer announcements we identified the topics in each of them. For the node 1191 there are 3 topics *i.e.*, it is new functionality, it reads announcements, and it will be launched in March. For the node 938 there are 4 topics *i.e.*, it is a festival, it will happen at SAMUCA, it is an invitation, and there will be food and gifts.

Table 2 presents two columns for each shorter announcement, one presents the total number of words provided by the participant, the other the number of correct words provided from the total number. This way we are able to evaluate the quality of the responses.

Considering that the objective for the longer announcements were to write the understanding of the media presentation instead of the exact words, it was considered the topics mentioned in the responses provided by the participants. We only listed the correct data; participants were not supposed to write exact words but their understanding.

Table 2. The responses provided by the participants from the activity of hearing and/or seeing announcements.

	Node 855		Node 974		Node 1191	Node 938
	Total	Correct	Total	Correct	Correct	Correct
Participant 1	3	2	2	0	0	2
Participant 2	4	4	5	5	2	3
Participant 3	7	7	5	5	2	4
Participant 4	6	5	5	5	3	3
Participant 5	8	3	2	2	1	2
Participant 6	4	3	3	3	2	3
Participant 7	7	7	4	4	3	4
Participant 8	5	5	3	2	1	3
Participant 9	7	7	5	4	1	4
Participant 10	7	7	3	3	1	3
Participant 11	5	5	3	2	2	3

In order to evaluate the results we considered the relative covering of the correct responses for each announcement, using the set of individual responses, the mean, mode and standard

deviation (see Table 3). Based on the mean we verified that, for the shorter announcements, participants were more precise using the audio version of the announcement. For the longer announcements occurred the opposite. As we can observe, the correctness rate among the participants vary significantly as for instance the contrast between Participant 1 and Participant 7. While Participant 1 was not able to understand the media of two announcements, Participant 7 hit most of the items.

Table 3. Analysis of the responses from the activity of hearing and/or seeing announcements.

	Node 855		Node 974		Node 1191	Node 938	
	Total (%)	Correct (%)	Total (%)	Correct (%)	Correct (%)	Correct (%)	Std. Dev. Participant
Participant 1	42.86	28.57	40.00	0.00	0.00	50.00	24.31
Participant 2	57.14	57.14	100.00	100.00	66.67	75.00	18.38
Participant 3	100.00	100.00	100.00	100.00	66.67	100.00	16.67
Participant 4	85.71	71.43	100.00	100.00	100.00	75.00	15.53
Participant 5	114.29	42.86	40.00	40.00	33.33	50.00	6.90
Participant 6	57.14	42.86	60.00	60.00	66.67	75.00	13.64
Participant 7	100.00	100.00	80.00	80.00	100.00	100.00	10.00
Participant 8	71.43	71.43	60.00	40.00	33.33	75.00	21.33
Participant 9	100.00	100.00	100.00	80.00	33.33	100.00	31.45
Participant 10	100.00	100.00	60.00	60.00	33.33	75.00	27.90
Participant 11	71.43	71.43	60.00	40.00	66.67	75.00	15.89
Mean -Node	81.82	71.43	72.73	63.64	54.55	77.27	
Mode -Node	100.00	100 and 71.43	60.00	100 and 40	100 and 33,33	75.00	
Std. Dev. -Node	23.12	26.34	24..12	32.02	30.81	17.52	

At the end, all participants answered a 5 point Likert Scale questionnaire. With this questionnaire, we wanted to confirm the results observed during the activity, asking them directly about their impressions on details of the presenter. Moreover, before the end of the workshop, a quick and informal final discussion reviewed their impressions on the activity of the day.

The questions were either related to the voice that read the announcement (Festival), the video with a human face (Virtual Presenter), or related to the integration of both voice and video instantiated at *Vila na Rede*. Table 4 presents results of the questionnaire. As for the results of the first task (transcription and understanding of audio and audio+video), they are detailed in Appendix 1, in Portuguese. These results are discussed in the next section.

Table 4. Results from the questionnaire.

Question	Answers	Total
About the spoken add, you understood:	Not a single word	0
	Almost nothing	2

	About half of the words	3*
	Almost all of the words	4
	Everything	2
The voice is:	Too slow	2
	Slow	2
	Adequate	2
	Fast	5
	Too fast	0
The quality of the voice is:	Very poor	0
	Poor	3
	Fair	7
	Good	1
	Very good	0
Are voice and presenter a good match?	No match at all	7
	Almost no match	1
	Matches somehow	3
	Good match	0
	Total match	0
The video is	Too artificial	4
	Artificial	3
	Not artificial nor natural	3
	Natural	1
	Very natural	0
Do the words seem to be said by the presenter?	Not at all	4
	A little	3
	More or less	3
	It is a match	1
	Yes	0
The video contributed to the comprehension of the audio from the announcement	Very little	2
	A little	1**
	Did not contribute but did not hinder	7
	Contributed	1
	Great contribution	0
The size of the video is	Terrible. Could be a lot bigger	0
	Not good, could be bigger.	2***
	Fair	7
	Good	2
	Excellent	0

*** The participant added the following comment: "except for the longer announcements"**
**** The participant added the following comment: "because I paid more attention to the voice"**
***** The participant crossed the "not good", and wrote "Good!"**

5. Discussion

Based on the data from the responses of the activity of hearing and/or seeing announcements content we observed that: 1) the more precise transcriptions came from the audio-only media, for shorter announcements; 2) the more complete understandings came from using Virtual Presenter media, for longer announcements; 3) and there wasa significant variance among the participants' individual results. In the transcription of the media content for shorter announcements, participants were slightly better when using audio-only content (71.43%) than using Virtual Presenter (63.64%). We believe that the objective of the activity contributed for this result; as the participants were really concerned about writing the text they just heard, they might not pay attention to the video. When considering the statistical mode of the correct responses we observe that both audio only and the virtual presenter obtained almost the same scores - *i.e.* 100% and 71.43% for audio-only, and 40% and 100% for Virtual Presenter.

When the activity consisted of writing the understanding of the media content of longer announcements, we observed a difference in the kinds of media. Virtual Presenter obtained a rate of 77.27% of correctness while audio-only reached 54.55%. By analyzing the statistical mode we reinforce the difference *i.e.* a draw between 33.33% and 66.77% for audio-only, and 75% for Virtual Presenter. Additionally, the standard deviation of the Virtual Presenter is about a half (17.52%) of the audio-only, which indicates an actual improvement in the understanding of the media..

A comparative analysis between audio-only and Virtual Presenter for each participant results reveals that: results vary significantly among the participants (*e.g.* average of 19.64% of correctness by Participant 1, and 95% by Participant 7). When considering the participants that obtained the best results (*i.e.* 3 and 7), we identify that one got better results from the audio-only media and the other from the Virtual Presenter. However, when computing the results of the rate of correctness of the audio-only in relation to Virtual Presenter media for each participant, we verified that for 9 (of 11) of them got better results using Virtual Presenter.

The answers from the questionnaire were grouped and summed for analysis and comparison with the statistical data. All answers were abstracted and scaled as a typical 5-point Likert item, so that the participants' opinions were considered as being: 1 - highly negative; 2 - negative; 3 - neutral; 4 - positive and 5 - highly positive. The results presented on Table 4 in the previous section were grouped and are shown on Table 5 and Figure 3.

Table 5. Results for voice, video and voice+video.

	minimum	Mode	Maximum
Voice			
Understanding	2	4	5
Speed	1	4	5
Quality	2	3	4
voice + video			

Match	1	1	4
Reality	1	1	3
Video			
Reality	1	1	4
Size	2	3	4
Contribution to comprehension	1	3	4

In general, it seems that the acceptance of the voice alone was good. The results from the questionnaire confirmed our initial hypothesis that the participants enjoyed synthesized voices. Most of the participants reported to be able to clearly understand the voice and the opinion about its quality ranged from fair to good. But, as accounted by the questionnaire's results, the voice did not make a good pair with the video. The characteristics resulted from this combination voice+video received mostly negative marks: the assembly was reported to be unnatural and voice and video did not match (not in the sense that mouthing and speech were not encompassed, but that the voice did not look like that it came from the person represented in the video).

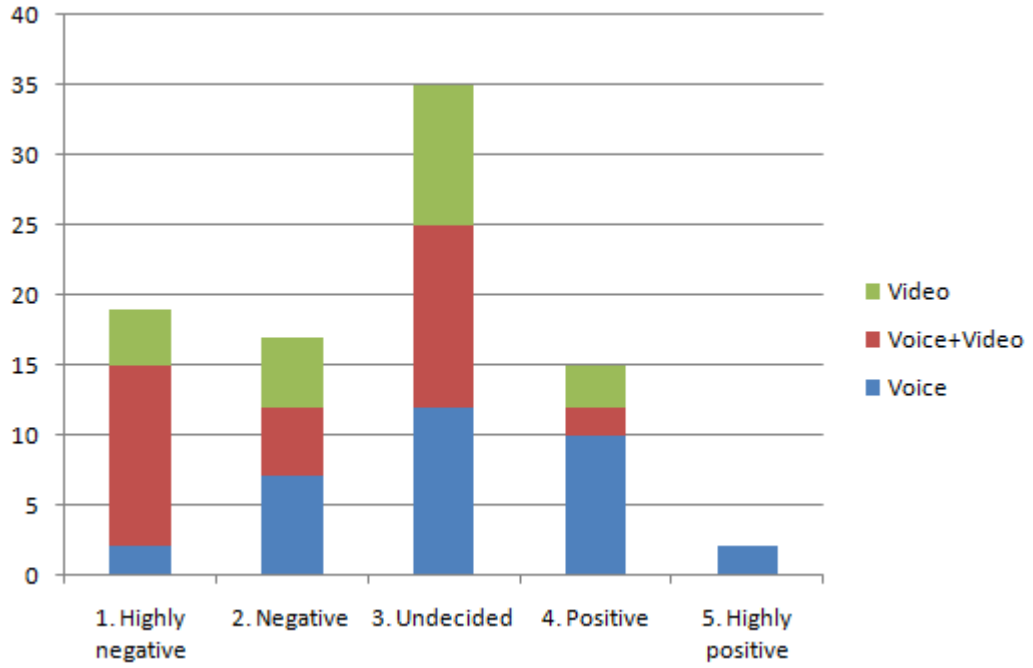

Figure 3. Distribution of answers from the Likert scale.

As the voice does not appear to match the face presented in the video, this seems to have caused a negative influence in the overall acceptance of the video. One factor that might also have impacted in the rating of the video is that most participants, as they were making intensive effort in the task of hearing and taking notes, did not pay much attention to the video. This was observed by researchers and also, it was added as a side note in the questionnaire, by one of the participants. Further investigation is needed to know how much influence it has on the Virtual Presenter's acceptance, the fact that the human face

exposed in the video does not show any emotions. Another point to be considered in future investigations is to check for results in more natural settings, i.e. situations where users are not demanded to provide written and accurate account from the audio.

The data obtained from the observations and analysis of the forms (level of comprehension) matches the impressions they orally reported to have about the feature;i.e. that the video had a positive influence in the understanding of the spoken announcements. However, the data collected from the questionnaire seem to go on the opposite direction. We believe that the way the answers from the questionnaire were written did not correctly reflect the real negative/positive scale as the one proposed by a 5-point Likert scale. For example, the question "the voice is:" had the choices: "1. too slow, 2. slow, 3. adequate, 4. fast and 5. too fast". In fact, the most positive answer would be "adequate", while "too fast" would be considered a rather negative one. Nonetheless, "adequate" as taken as a neutral response (2 participants' choices), and "too fast" was considered as highly positive (no participant choice). This should not invalidate our experiment - as we had redundant sources of collected material to base on - but it confirms the need for further investigation.

6. Conclusion

The Virtual Presenter - a talking head that has been adjusted for our context- has been incorporated into the Inclusive Social Network *Vila na Rede*, aiming at helping users to make better use of the information available in the system. A 2D facial animation that realistically reproduces speech articulatory movements was implemented at *Vila na Rede*, together with a female-converted voice from the Festival text-to-speech tool.

This paper described the process in which the Virtual Presented was incorporated into our system and it describes the activities that evaluated this mechanism. The results indicate that the video with a human presenter moving her lips accordingly might help users in the understanding of longer audio extracts. However, it is important to have voice and face in harmony, providing a more natural aspect to the presenter.

The experiment contributed to clarify our target users preferences and use of this solution regarding a virtual presenter in the system, but further investigations are needed in order to obtain more conclusive results about the use of the Virtual Presenter at *Vila na Rede*.

Acknowledgments: This work is funded by Microsoft Research – FAPESP Institute for IT Research (#07/54564-1), and partially by FAPESP (#07/02161-0) and CAPES (#01-8503/2008). The authors also thank Professor José Martino and Paula Costa; colleagues from IC/UNICAMP, NIED/UNICAMP, InterHAD and CenPRA for insightful discussion.

Referências

Bohus, D.; Horvitz, E. 2009. Dialog in the Open World: Platform and Applications. In: Proceedings of ICMI-MLMI'09. Cambridge, MA. p. 31-39.

Cao, Y.; Tien, W.C.; Faloutsos, P., Pighin, F. 2005. Expressive Speech-Driven Facial Animation. In: ACM Transactions on Graphics (TOG), Col. 24, issue 4, p. 1283-1302.

Costa, P. D. P. 2009. Animação facial 2D sincronizada com a fala baseada em imagens de visemas dependentes do contexto fonético. Mastering thesis of the School of Electrical and Computer Engineering at University of Campinas, Brazil.

DiPaola, S.; Arya, A. 2007. A Framework for Socially Communicative Faces for Game and Interactive Learning Applications. In: Proceedings of the 2007 conference on Future Play. Toronto, Canada. p. 129-136.

Ezzat, T.; Poggio, T. 1998. MikeTalk: A Talking Facial Display Based on Morphing Visemes. In: IEEE Computer Animation, Center for Biological & Computational Learning and the Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Philadelphia, CA, EUA.

FFmpeg. 2010. http://www.ffmpeg.org/.

Hayashi, E. C. S.; Baranauskas, M. C. C. 2009. Communication and expression in social networks: Getting the "making common" from people. In Proceedings of the 2009 Latin American Web Congress, Joint LA-WEB/CLIHC Conference, Mérida, Mexico. IEEE Computer Society 2009, pp. 131-137.

Hayashi, E. C. S.; Baranauskas, M. C. C. 2008. Facing the digital divide in a participatory way – an exploratory study. In 1st IFIP Human-Computer Interaction Symposium (HCIS 2008), IFIP World Computer Congress (WCC 2008). Boston: Springer, v. 272. p. 143-154.

Long Tail Ad Solutions. 2010. JW Player. http://www.longtailvideo.com/players/jw-flv-player.

Mumford, E. 1964. *Living with a computer*. London, England: Institute of Personnel Management.

Neris, V. P. de A.; Almeida, L. D. A.; de Miranda, L. C.; Hayashi, E. C. S.; Baranauskas, M. C. C. 2009. Towards a Socially-constructed Meaning for Inclusive Social Network Systems. In: ICISO 2009, Beijing-China. p. 247-254.

Open Computer Vision Library. 2010. http://sourceforge.net/projects/opencvlibrary/.

Parke, F. I. 1972. Computer generated animation of faces. In: Proceedings of the ACM annual conference SIGGRAPH. Boston, p. 452-457.

Salvi, G.; Beskow, J.; Al Moubayed, S. and Granström, B. 2009. SynFace—Speech-Driven Facial Animation for Virtual Speech-Reading Support. In: EURASIP Journal on Audio, Speech, and Music Processing Volume 2009, p. 10.

Stamper, R. K.; Althans, K.; Backhouse, J. 1988. Measur: Method For Eliciting, Analysing and Specifying User Requirements. In: Computerized Assistance During the Information Systems Life Cycle. North-Holland, p. 67-115.

University of Edinburgh. 2010. The Centre for Speech Technology Research - The Festival Speech Synthesis System. http://www.cstr.ed.ac.uk/projects/festival/.

Appendix A

Transcriptions (in Portuguese)

Group A

1)

1 - "ponto cruz. anunciar"

2 - "fuxico ... fuxico"

3 - "Terá uma apresentação em março, no Vila na Rede virtual"

4 - "Festa junina com muita festa e brinde para todos"

2)

1 - "bordados - fitas - ponto cruz"

2 - "boneca de fuxico, boneca feita de fuxico"

3 - "lançamento do apresentador virtual em março:"

4 - "anúncio da festa junina do Samuca:"

3)

1 - "bordados fitas, ponto cruz e outros preços a combinar"

2 - "bonecas de fuxico. Boneca feita de fuxico"

3 - "Nova funcionalidade apresentador virtual.. lançamento em março"

4 - "festa junina no Samuca. Vai ter brindes, bancas típicas. Venham prestigiar no arriá do Samuca."

4)

1 - "Bordados fita ponto cruz e outros combinar"

2 - "Boneca de fuxico ... ? essa é uma boneca de fuxico"

3 - "Novas funcionalidades do vila: apresentador virtual e apresentadora virtual que lê os anúncios. Lançamento em março."

4 - "Anúncio do arraial do Samuca com comidas e bebidas e sorteio de brindes"

5)

1 - "bordado e ponto cruz e outros. Bordado cruz e ponto e outro"

2 - "boneca de fuxico"

3 - "março lançamento no Vila na Rede que será o novo lançado em março"

4 - "a festa junina venha participar vai ter barraca típica"

Group B

1)

1 - "fuchico. Boneca fuchico"

2 - "ponto. bordados pinta combinar"

3 - "<não entendi a letra> venha participar do grande arrraiá do Samuca. teremos muito para todos"

4 - "Nova funcionalidade no Vila na Rede. Apresentador virtual, teremos a apresetadora virtual. terão novos anúncios. O lançamento será em março."

2)

1 - "Boneca de fuxico e... feita de fuxico"

2 - "Bordados, fita, ponto cruz e outros. Preços a combinar"

3 - "Festa junina Samuca, venha prestigiar o arraiá do Samuca, teremos muitos brindes, venha nos visitar"

4 - "Nova funcionalidade no Vila. Teremos a apresentadora virtual no Vila que lerá os anúncios para você. O lançamento será em março"

3)

1 - "boneca fuxico e ponto"

2 - "Bordo em fitas e outros preços a combinar"

3 - "Grande festas juninas venha nos prestigiar. Haverá barracas típicas e outros"

4 - "Anúncios do Vila na Rede erá nova virtual que será agora em março"

4)

1 - "boneca de fuxico ensaiado uma boneca feita de fuxico"

2 - "bordado fitas pontos cruz e outros preços a combinar"

3 - "festa junina venham prestigiar. Está convidando para festa junina do Samuca"

4 - "Novo apresentador virtual no Vila na Rede"

5)

1 - "Boneca feita de fuxico"

2 - "bordados fita ponto cruz e outros preços a combinar"

3 - "Festa junina no Samuca venha prestigiar e participar teremos na festa no Samuca"

4 - "Nova fucionanlidade no vila apresentador virtual que anuncia em março"

6)

1 - "Meu ... --- ka fuchico --- fuxico"

2 - "os pontos fita cruz os preços combinar"

3 - "É para participar de uma grande festa junina. Terá prendas - pipoca amendoim. Prestigiar . Não perca"

4 - "apresentador virtual. Vai acontecer na Vila União grande lançamento em março"