# Learning Icelandic language and culture in Virtual Reykjavik: starting to talk

Branislav Bédi[1], Birna Arnbjörnsdóttir[2],
Hannes Högni Vilhjálmsson[3], Hafdís Erla Helgadóttir[4],
Stefán Ólafsson[5], and Elías Björgvinsson[6]

**Abstract**. This paper describes how beginners of Icelandic as a foreign and second language responded to playing the first scene in Virtual Reykjavik, a video game-like environment where learners interact with virtual characters – Embodied Conversational Agents (ECAs). This game enables learners to practice speaking and listening skills, to learn about the language and culture by solving various tasks. A mixed-method pilot study examined how six learners responded. The results were divided into how learners perceived (1) playing the game in general; (2) interaction with virtual characters; (3) multimodal behaviour of the virtual characters, in particular the speech, facial expressions, hand gestures and body movements; and (4) what the game helped them learn. In conclusion, some learners reported they learned to start speaking in Icelandic and that the spoken behaviour of virtual characters resembled real Icelanders.

**Keywords**: computer games, embodied conversational agents, multimodal behaviour, natural language, virtual learning experience.

## 1.    Introduction

Virtual Reykjavik is an online language and culture training application (Vilhjálmsson, 2011) designed for beginning adult learners of Icelandic as a foreign

1. University of Iceland, Reykjavík, Iceland; brb19@hi.is
2. University of Iceland, Reykjavík, Iceland; birnaarn@hi.is
3. Reykjavik University, Reykjavík, Iceland; hannes@ru.is
4. Reykjavik University, Reykjavík, Iceland; hafdis13@ru.is
5. Reykjavik University, Reykjavík, Iceland; stefanola13@ru.is – Northeastern University, Boston, MA, United States; stefanolafs@ccs.neu.edu
6. Reykjavik University, Reykjavík, Iceland; eliasb13@ru.is

or second language. This video game-like environment enables learners to practice spoken language by talking to virtual characters and listening skills by listening to the virtual characters' answers. Conversational interactions are a good source for authentic language input (Krashen, 1982) and output (Swain, 2007) with potentials of developing learner's language skills (Ellis, 1991). Speaking is realised by means of Google speech recognition service for Icelandic; the speech is automatically transcribed for learners to read what has been said in a lower text window. Character responses are synthesised using the Ivona Text To Speech (TTS) system for Icelandic.

The current version uses game-based learning (Meyer, 2009), task-based learning (Ellis, 2003), and a communicative approach (Richards, 2006), supporting learning of new vocabulary through text windows of cultural information and transcribed dialogues. The main menu is in Icelandic, where learners can hover over words with a mouse cursor and see their English translation. Learners can find cultural information about selected buildings and monuments in a pop-up window when in pause or 'freeze' mode.

Figure 1.   Approaching strangers in Virtual Reykjavik



The virtual characters are designed to look and act lifelike (Prendinger & Ishuzuka, 2004). In this context, they are also called ECAs because of their embodiment within a virtual environment, taking human form, and knowing how to use their body in conversations (Cassell, 2000). They are able to "grasp situations, react to contingencies and look aware of their surroundings" (Pedica & Vilhjálmsson, 2012, p. 1). Their social and conversational behaviour is multimodal and includes speech, facial expressions, hand gestures, head, and body movement. They are designed to look natural based on research into real-life behaviour. In this way, learners can get familiar with the social behaviour and learn about culture and language remotely.

In that respect, Prada and Paiva (2014) suggest that it is important to study human social interaction in order to help create a highly interactive exchange between human users and virtual agents; the effects are stronger if the artefacts, i.e. agents, present autonomous and proactive behaviour. About 144 video recordings capturing first encounters between native and non-native speakers in the same conversational situation as in this game were collected in downtown Reykjavik and analysed. This multimodal corpus provided authentic data for a discourse modelling. Specific communicative functions and behaviours associated with those functions were implemented into the architecture of the system that the virtual characters use to generate behaviour, allowing for a more realistic human-to-agent interaction.

In the first scenario, currently including only one scene, *Týnda hljómsveitin* [The Lost Band], and one chapter, *Hvar er Hitt Húsið?* [Where is Hitt Húsið?], learners have three tasks to solve: (1) approach a virtual character, who is a native speaker of Icelandic and get his/her attention by greeting appropriately (Figure 1); (2) ask for directions to a particular place; and (3) say goodbye.

## 2.      Method

A mixed-method approach consisting of a questionnaire, an interview and a video recording of each participant while playing was used. The pilot study took place in a laboratory setting. Participation was anonymous and informed consent was obtained. Six learners of six different nationalities participated; four female and two male, aged 22-31; three were temporary and three permanent residents. Five were beginners and one a false beginner (intermediate); their level of Icelandic was A1 on the Common European Framework of Reference for languages.

The objectives focused on examining how learners perceived: (1) playing the game in general; (2) the interaction with virtual characters that possess communicative skills based on conversations between people in the real-life situations, because interacting with such agents may promote communication skills (Johnson et al., 2004) and facilitate learning (Shaw, Johnson, & Rajaram, 1999); (3) the multimodal behaviour of the virtual characters; and (4) what the game helped them learn.

## 3.      Learner's perception and learning effect

The results indicated that half of the learners found playing the game enjoyable and exciting, while two of six described it as frustrating and boring, and one as neither.

The enjoyment and excitement referred to the learners being able to talk to virtual characters and see what they were going to say. Frustration was mostly caused by the speech recognition not handling improper pronunciation of certain consonants, e.g. <ʁ> instead of <R>, and due to speaking too softly or fast into the microphone, but also by its lagging; moreover also by the lack of information about the goals of the game and not being able to ask questions other than those given in the tasks. Learners perceived boredom mostly due to the slow movement of the avatar. Two of six learners found the game easy and four as neither easy nor difficult. None of them found the game difficult.

Regarding the specific multimodal behaviour (Figure 2), spoken language was perceived as the most natural and variation was important: "[I]t wasn't just one *góðan daginn* [good day, with definite article], but also *góðan dag* [good day, without definite article] as well, so it seemed like you were actually speaking to someone". However, it was pointed out that in some cases the virtual characters did not respond adequately to all forms of greetings, such as *blessuð* [greeting a female], *blessaður* [greeting a male], etc. The reason is due to lack of informal greetings. Two learners described the facial expressions as robotic. Some learners also described the body movement as spontaneous. Hand gestures were registered only during pointing when virtual characters were giving directions or crossing their arms.

Figure 2.  Perception of specific multimodal behaviour
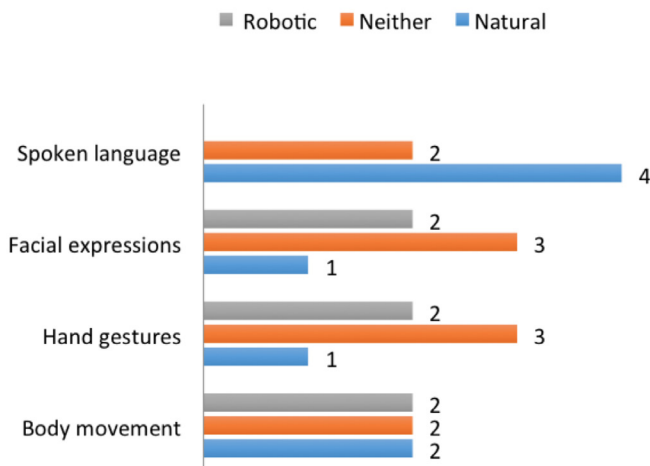


Regarding the overall multimodal behaviour (Figure 3), some learners commented that when speaking to virtual characters they had difficulties being understood;

also some virtual characters when standing on their own looked like "I-don't-wanna-talk-to-you kind of". In addition, there was a slow pace of moving around; the interaction was sometimes distracted by focusing on the red speaking button, which signalised when the learner could speak; the interaction felt natural by half, but on the other hand cold by more than a half of learners mainly due to the lack of emotion expressed by the characters.

Figure 3.  Perception of the overall behaviour
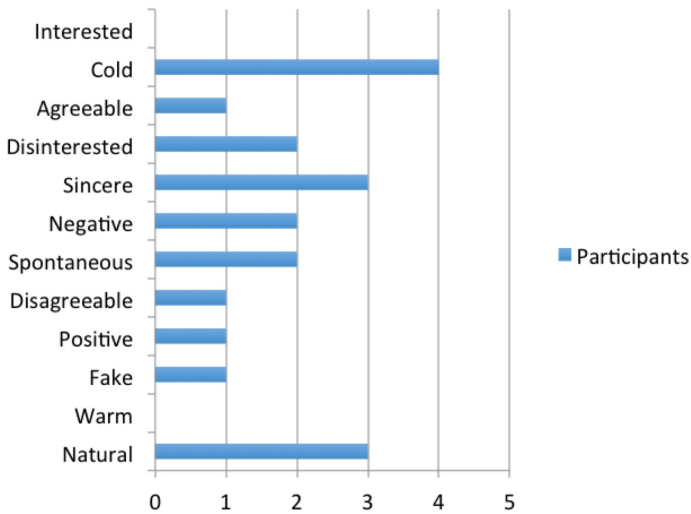


The game had a reported educational effect on five out of six learners; learning about famous places in downtown Reykjavik and to start talking in Icelandic. One learner found it pointless because of pronunciation issues and experiencing difficulty being 'not understood'. Overall, four learners stated that once the game is ready, it would help others learn about buildings and monuments and make it easier to start speaking and communicating in Icelandic. They felt that a virtual environment would be easier than learning in reality with real people face-to-face.

## 4.    Conclusion

Five out of six learners found the game educational, of which three learners reported they learned (1) spoken language skills when initiating a conversation, and (2) about cultural sights and buildings in downtown Reykjavik. A gap

between a passive and an active use of language can be bridged when learners get to practice the language in a conversation with virtual agents. Quoting from one learner: "I think it's very helpful, like when [the game] would be done, […] especially in practicing conversations, because when you learn a language the hardest part is to start speaking". In this way, the game may help facilitate face-to-face communication in reality. The game's ultimate goal is to promote learning of language and culture based on real language use with virtual characters acting like native speakers. However, the multimodal behaviour needs to be fine-tuned yet. Adding a smile at the end of a conversation, which can also be observed in recordings from real life, can make the virtual characters friendlier. It is, nonetheless, difficult to draw a general conclusion on whether or not the design of virtual characters can have a similar effect on learners in another computer game. Since the aim of the Virtual Reykjavik project is to create lifelike virtual characters with natural multimodal behaviours, more focus must be placed on incorporating those into the system, and more research needs to be done in real-language data to reach perfection in modelling believable embodied conversational agents for language learning games.

## References

Cassell, J. (2000). Nudge, nudge, wink, wink: elements of face-to-face conversation for embodied conversational agents. In J. Cassell, J. Sullivan, S. Prevost, & E. Churchill (Eds), *Embodied conversational agents* (pp. 1-27). Cambridge, MA: MIT Press.

Ellis, R. (1991). The interaction hypothesis: critical evaluation. *Paper presented at the Regional Language Seminar, Singapore, March 1991*. Eric – US Department of Education. http://files.eric.ed.gov/fulltext/ED338037.pdf

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.

Johnson, W. L., Choi, S., Marsella, S., Mote, N., Narayanan, S., & Vilhjálmsson, H. (2004). Tactical language training system: supporting the rapid acquisition of foreign language and cultural skills. In *Proceedings of InSTIL/ICALL - NLP and Speech Technologies in Advanced Language Learning Systems, June 17-19, Venice*.

Krashen, S. D. (1982). *Principles and practice in second language acquisition.* Pergamon Press.

Meyer, B. (2009). Designing serious games for foreign language education in a global perspective. In A. Méndez-Vilas, J. Mesa González, J. A. Mesa González & A. Solano Martín (Eds), *Research, reflections and innovations in integrating ICT in education* (Vol. 2, pp. 715-719).

Pedica, C., & Vilhjálmsson, H. H. (2012). Lifelike virtual characters using behavior trees for social territorial intelligence (Demo/Poster). In *Proceedings of ACM SIGGRAPH'12, Los Angeles, August 5-9*. http://www.ru.is/kennarar/hannes/publications/SIGGRAPH2012_poster.pdf

Prada, R., & Paiva, A. (2014). Human-agent interaction: challenges for bringing humans and agents together. In *HAIDM – 3*rd *International Workshop on Human-Agent Interaction Design and Models held at AAMAS'2014 - 13*th *International Conference on Autonomous Agents and Multi-Agent Systems*. Paris, France.

Prendinger, H., & Ishuzuka, M. (Eds). (2004). Introducing the cast for social computing: life-like Characters. In ). H. Prendinger & M. Ishuzuka (Eds), *Life-like characters: tools, affective functions, and applications* (pp. 3-16). New York, NY: Springer..

Richards, J. C. (2006). *Communicative language teaching today.* Cambridge University Press.

Shaw, E., Johnson, L., & Rajaram, G. (1999). Pedagogical agents on the Web. In *Proceedings of the third annual conference on Autonomous Agents, AGENTS'99* (pp. 283-290). New York, NY: ACM.

Swain, M. (2007). Fálagstilgátan: Kenningar og Rannsóknir. In A. Hauksdóttir & B. Arnbjörnsdóttir (Eds), *Mál málanna* (pp. 117-136). Reykjavik: Stofnun Vigdísar Finnbogadóttur í erlendum tungumálum.

Vilhjálmsson, H. H. (2011). Icelandic language and culture training in Virtual Reykjavik. The Icelandic Research Fund 2012. Project Grant - New proposal (pp. 1-13).

**CALL communities and culture – short papers from EUROCALL 2016**
**Edited by Salomi Papadima-Sophocleous, Linda Bradley, and Sylvie Thouësny**