Was it something I said? Facial Expressions in Language Learning

John Sloan, Julie Carson-Berndsen

ADAPT Centre, School of Computer Science, University College Dublin, Ireland

`john.sloan.1@ucdconnect.ie, julie.berndsen@ucd.ie`

Figure 1: *Sequence of change in expression of the Mery [1] avatar after each of the participants' 6 input sentences.*

Abstract

This paper describes an experiment to evaluate facial expressions of an animated avatar as a means of providing feedback to non-native English learners on language production. The aim of the study was to ascertain whether native and non-native English speakers interpret and respond differently to facial expressions and whether such expressions have a role to play as feedback in emerging language learning technologies. Native English speakers and non-native English learners took part in an experiment where facial expressions were presented as a response to their textual input sentences and were asked for their interpretation of the change in expression (or otherwise). Furthermore, it was investigated to what extent non-native learners subsequently altered their language behaviour in line with their perceived interpretation of the expression. The majority of non-native learners attributed a change in facial expression, where the avatar looked away, to errors in language production in the preceding sentence and they reduced the syntactic complexity of the following sentence accordingly. The results underpin the potential of facial expressions as a feedback mechanism for language learning and the insights will now be deployed in an effective and engaging personalised e-learning language platform.

Index Terms: second language acquisition, feedback, facial expressions, avatars

1. Introduction

Corrective feedback in second language acquisition (SLA) has been shown to improve learners' retention of vocabulary and grammar [2]. Recasts, clarification requests and comprehension checks are among the verbal feedback forms which have proven most effective [3,4], however, there has been no real consensus on which of these mechanisms produces the optimal learning outcome. In addition, the impact of non-verbal, paralinguistic feedback has received very little attention [2]. With the growth of Computer Assisted Language Learning (CALL), the question of when and how to provide feedback to learners is of immediate concern, especially when the effectiveness of current CALL systems is limited [5]. While e-learning can provide learners with an increase in content, interactions and a variety of methods of study, the distance between learners and instructors (or indeed other learners) can lead to a shallow experience with inaccurate, ineffective or even a lack of feedback.

One area of CALL which has received a lot of recent interest is the use of robots which interact with learners in the same physical space. These robots can be controlled remotely [6] or automatically [7]. Advantages offered by this form of interaction include increased engagement and a richer educational experience. Improvements in performance on tasks involving puzzle-solving in adults [8] and prime number learning in children [9] have been demonstrated. However, the body of research in this area is still small and is based on short-term studies. Furthermore, several negative effects have also been identified. These include deception, privacy issues, lack of accountability and, particularly in the case of children, attachment to the robot, loss of human contact and impeded social skill development [9,10].

In contrast to physically embodied robots, screen technologies are widely used to deliver CALL content. The recent global proliferation of internet and mobile technology (43% own a smartphone, 76% report using the internet [11]) means through the screen, CALL is readily accessible. Over the past decade, the most popular language learning applications have reached tens of millions of registered users [12]. For SLA, screen-based technologies offer advantages over other interactive media due to the privacy afforded to the learner in a task where making errors is common. Success in acquiring a second language requires not only considerable time and effort, but also the ability to learn from these errors. However, Cross Linguistic Interference (CLI) can lead to the repeated production of 'basic' mistakes which often results in anxiety, as the learner fears appearing foolish in front of peers [13]. In the early stages of language learning, anxiety can be low, but may grow through negative learning experiences emanating from negative perceptions of others [14,15]. Educational gains from short-term interaction with robots might not continue in the long-term if users' frequent errors are salient to others in the shared physical space.

Using facial expressions as the means of providing feedback through a screen offers an opportunity to reach large numbers of users while increasing engagement and pedagogical effectiveness. Current, popular screen-based

CALL applications (e.g. 'Duolingo', 'Babbel') typically provide feedback through simple video/audio signals which are intuitively understood by the user: green vs. red colours, a progress bar moving forward vs. backward, pleasing ping sounds vs. harsh klaxons. Facial expressions offer a potentially more powerful medium to provide feedback, as many expressions can be intuitively understood [16] and we have an innate preference to pay attention to faces and face-like stimuli [17]. Additionally, recent human-computer interaction studies have shown our lexical and gestural alignment with human-like avatars is comparable to that with real humans [18,19]. Through the medium of a screen, which can be kept relatively private, and the inaudible nature of facial expressions, the user is afforded a sense of anonymity [20] while engaged in the interaction. These factors could aid retention rates in language learning applications where dropout rates are high, e.g. 54% use Babbel for less than a month [21].

This paper presents an experiment to evaluate the use of an avatar's facial expressions as a potential feedback mechanism for language learning. Based on first-hand experience of the first author teaching Korean-learners of English for many years, a personalised e-learning platform is under development which aims to enhance non-native learner engagement and performance, and overcome anxiety issues associated with errors. The aim is to create an environment where language learners can practise producing samples of language while receiving immediate, accurate feedback which serves the function of guiding the learner to produce well-formed utterances. The platform currently uses the Mery avatar [1] with animations which have been created specifically for the purpose of providing expressive feedback. The animations run in modern web browsers at 10fps in 400x320px resolution. This produces smooth, realistic animations with the result akin to an animated movie. The avatar was chosen in preference to a life-like model due to unnatural movements of life-like models causing feelings of unease in users [22], the perceived intensity of expressions in animated avatars are higher than life-like models [23], and the avatar offers more flexibility in the creation of animations.

The remainder of this paper is structured as follows. Section 2 focusses on the facial expressions presented to the participants and the hypotheses being tested. The experiment to evaluate the facial expressions is described in section 3. Section 4 presents the results, with the discussion following in Section 5. Finally, Section 6 draws some conclusions and outlines future work.

2. Facial Expression Feedback Hypothesis

The research presented in this paper represents the first step in the investigation of the interpretation and response of non-native learners (NNS) and native speakers (NS) to changes in facial expression of an animated avatar. The sequence of expressions (e1 to e6), as depicted in Figure 1, are the respective responses to six sentences (s1 to s6) input by the user to the system. The focus will be on reactions to one specific, clear and distinct change in expression - from smiling and looking straight ahead to slightly frowning and looking down and to the side (e2 → e3 in Figure 1). The change from eye-contact to averting eye-contact is based on task-oriented studies where gaze aversion has been shown to correlate with the cognitive difficulty of the task [24,25]. Therefore, the

avatar looking away following an input sentence indicates an increased cognitive load.

The broad hypothesis to be tested is drawn from observances of NNS in interaction with NS, namely:

There is a difference between NS and NNS in how they interpret and respond to facial expressions of a NS interlocutor.

NNS, particularly proficient speakers, often appear to reduce the length and complexity of utterances when met with a NS facial expression which would be typically described as confused. This may be due to a recognition by the NNS that their production is causing additional cognitive effort on the part of the NS which must be reduced.

In the experiment presented in the next section, the animated avatar was used in place of a NS interlocutor and participants were informed that the avatar could read and understand English sentences. The broad hypothesis can then be refined to refer to facial expressions of this avatar. This hypothesis was broken down into four sub-hypotheses.

*   **H1**: There is a difference between NS and NNS in the emotion attributed to the avatar's e2 → e3 change in expression.

*   **H2**: There is a difference between NS and NNS in the reason given for this change in expression.

*   **H3**: There is a difference between NS and NNS in the complexity of sentence produced following an e2 → e3 change in expression.

*   **H4**: In the complexity of sentence produced following an e2 → e3 change in expression, there is a difference between NNS who are sensitive to small changes in the avatar's expression and NNS who are not.

The experiment to test these sub-hypotheses is now described.

3. Experiment

Two groups were selected for the experiment: native English speakers, and intermediate and above non-native learners of English. Participants (*N*=57) were gathered through postings on internet cafes for English language learners and snowball sampling. The experiment consisted of 2 parts: questions on emotions and English, and typing sentences to the avatar. All participants were aged 18+ and successfully passed part 1 of the experiment; this evidenced accurate recognition of images of human facial expressions and a command of English sufficient to partake in part 2. The sample contained 36 NNS and 21 NS, with participants indicating their nationality and L1 prior to beginning the experiment. Five participants were excluded from analysis: 2 incorrectly responded to questions in part 1 and 3 did not fully complete the sentences in part 2.

3.1. Experiment Design

A web application was developed for the experiment and hosted at emotionandlanguage.ucd.ie. Upon signing up for the research, participants were sent to the website where they were asked to complete a series of tasks. These tasks involved answering multiple choice questions, typing sentences and watching the avatar's subsequent expressions. These tasks were split into 2 parts which each took approximately 5-10 minutes to complete. Participants completed the experiment at a time and place of their choosing.

3.1.1. Part 1: Questions on English and emotions

Participants answered 15 multiple choice questions before interacting with the avatar. The first 10 showed a photograph of a person expressing an emotion with a choice of 4 words to describe the image. Participants chose one of the 4 options.
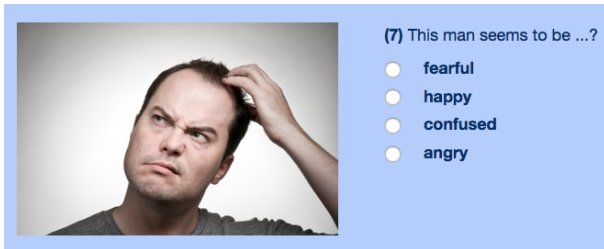

Figure 2: *Question on facial expressions and emotion.*

The purpose of the 10 expression questions was threefold: draw participants' attention to facial expressions, ensure NNS participants understood the vocabulary used to describe expressions and emotions, and confirm that participants gave typical responses to facial expressions. Some of the expression images were shown alongside only one word which matched (e.g. a photo of a smiling girl with the choices: ['angry', 'happy', 'sad', 'tired']), while others contained multiple (e.g. a man frowning and rubbing his eyes with the choices: ['excited', 'tired', 'sad', 'worried']). Five further questions on vocabulary and grammar were then presented. These ensured that NNS participants understood the English needed to give reliable answers to the questions in part 2.

3.1.2. Part 2: Sentences and Expressions

Our 4 sub-hypotheses were tested in part 2. Participants were introduced to the avatar, and after a short practice session, entered 6 sentences, one-by-one, about their favourite holiday. Regardless of linguistic content, the avatar changed expression after each sentence in the sequence shown in Figure 1. Participants were given no information as to the reason for the avatar's expression changes. Our focus was on the participants' response to the avatar's change (e2 → e3) after they entered sentence s3.

Figure 3: *View of the main interface where participants input sentences and monitor the avatar's expression (e2).*

For each sentence input, participants were shown the avatar, the prompt and a text box with cursor focus where they could type (Figure 3). Two seconds after a sentence was entered, the avatar began to move its head and eyes to simulate reading (Figure 4). Note that the movements of the avatar between all of the expressions are smooth.

Figure 4: *View of the avatar 'reading' s3.*

After 'reading' the sentence, the avatar either returned to looking straight ahead, or, in the case of interest after s3, looked down and to the left (Figure 5). Depending on the length of the input sentence, each animation lasted between 7 and 12 seconds. The facial expression changed in five of the six movements, with only the final sentence (s6) eliciting no change.

Figure 5: *View of the e3 expression after 'reading' s3.*

In this case, after the avatar 'read' s3 and changed expression, participants were asked to answer two questions: the first (**H1**) involved describing the avatar's emotion, and the second (**H2**), the reason behind this change of expression. The questions appeared 4 seconds after the change in expression and covered the avatar and the typed sentence on the screen as depicted in Figure 6.
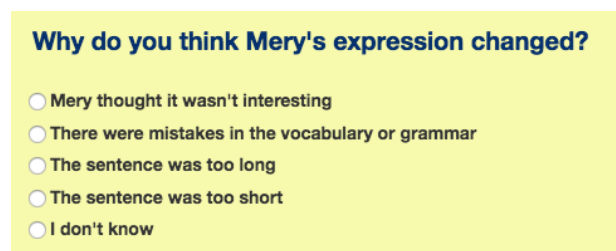

Figure 6: *Question box for H2 which appeared after the e2 → e3 expression change.*

3.1.3. Change in complexity

H3 and **H4** were tested by comparing the complexity of sentences produced immediately before and after e3 → e4. This involved scoring the complexity of s3 and s4 for both NS and NNS. While measures of syntactic complexity as development markers in L1 children, such as Developmental Sentence Scoring [26] and Index of Productive Syntax [27], and the extension of such systems to measure L2 proficiency [28,29] have been developed, there is no single reliable and valid measure for both NS and NNS complexity. In addition,

the NNS in this experiment came from a variety of countries and native languages (L1), which compounds the problem, e.g. for an NNS of a language which shares the head-first structure of English, a relative clause may be considerably easier to form than for an NNS whose native language is head-last.

With this in mind, it is still possible to formulate a general score of complexity (c-score) which correlates with human judgement. The formula used for this experiment is based on a small number of widely used factors which have been identified as useful markers in the literature and recent research on automatic syntactic complexity analysis [30]. These are: length of sentence, number of t-units (main clause with attached subordinate clauses [31]) per sentence, amount of subordination and sophistication of Noun Phrases (NP). For this experiment, the complexity score of each sentence was calculated using the following formula:

$$\text{c-score} = \text{tuM}(L / 3 + 2sC + cNP) \qquad (1)$$

where

- L = length of sentence in words
- sC = number of subordinate clauses
- cNP = number of complex NP (NP including head noun and an adjective or prepositional phrase)
- tuM = t-unit multiplier $(1 - \ln(tu / 3))$ where tu is the number of t-units

The weightings of each variable were determined based on theoretical ideals with testing and evaluation on sentences produced by NNS from a previous study. Therefore, subordinate clauses, which are often difficult to accurately produce for NNS were assigned a higher weighting than complex noun phrases. The additional length from higher t-units per sentence was balanced through 'tuM', which decreases exponentially from 1 in the order [1, 0.77, 0.63, 0.54 …]. An example calculation on 2 contrived sentences is shown below (sC is indicated with '{}', cNP with '_', and t-units are split with '|'):

s3: *"{Because <u>heavy rain</u> was falling everyday}, my sister, {who hadn't brought a raincoat}, didn't enjoy the holiday."*
s3 c-score: $1*(18/3 + 2*2 + \underline{1}) = \mathbf{11}$

s4: *"We stayed inside and didn't travel, | but she went out on <u>the third day</u> | and we had some fun."*
s4 c-score: $0.63*(19/3 + 0 + 1) = \mathbf{4.62}$

s3-s4 change in c-score: $4.62-11 = \mathbf{-6.38}$

To assess the validity of the c-score metric, all s3 and s4 sentences were independently scored for complexity by an expert with an MA in Linguistics and 10+ years experience teaching ESL. The scorer was informed of the factors taken into consideration for complexity, but not told of the weighting or method of calculation. The c-scores and the independent annotator's scores were normally distributed and exhibited a strong correlation $r(55)=.59$, $p<.001$.

3.1.4. Sensitivity

H4 investigates two populations of English learners: those with high levels of attention or sensitivity to facial expressions, and those with lower levels. We split our sample of English learners into 'sensitive' (N=11) and 'not sensitive' (N=25) groups based on the accuracy with which they recognised the avatar's changes of expression e4 → e5 and e5 → e6. e4 → e5 displays a small change in expression, with a slight lowering of the edges of the mouth and eyebrows; e5 and e6 exhibit the exact same expression. Only participants who correctly identified both changes were assigned to the 'sensitive' group. The change in s3 → s4 complexity between the groups was measured as an independent variable, 'Sensitivity'.

4. Results

In this section, results for each of the sub-hypotheses (**H1** to **H4**) are presented in turn.

H1: Change in expression

A Fisher's exact test [32] was performed to examine the relationship between NS and NNS in their identification of emotion in the avatar's expression after s3. No difference was found between the groups (p=.6). Both rated the expression of the avatar most frequently as 'confused' followed by 'thoughtful' as depicted in Figure 7.
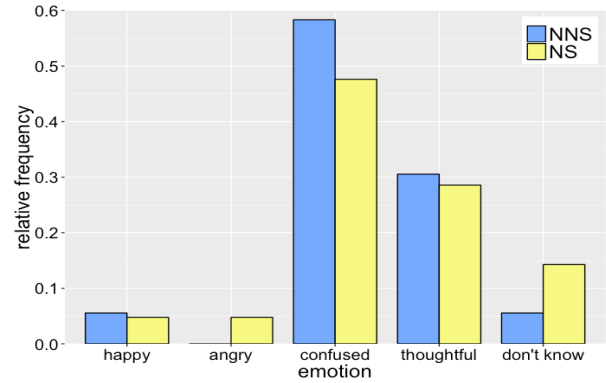

Figure 7: *Identification of emotion in expression e2 → e3.*

H2: Reason for change in expression

A significant difference was found in the reasons given for the avatar's change in expression after s3 (p=.002). Figure 8 shows the majority of NNS (53.8%) attributed the change to their mistakes in vocabulary or grammar in s3. In contrast, NS were unsure of the reasons for the expression change, selecting 'don't know' more frequently than NNS (47.6%).
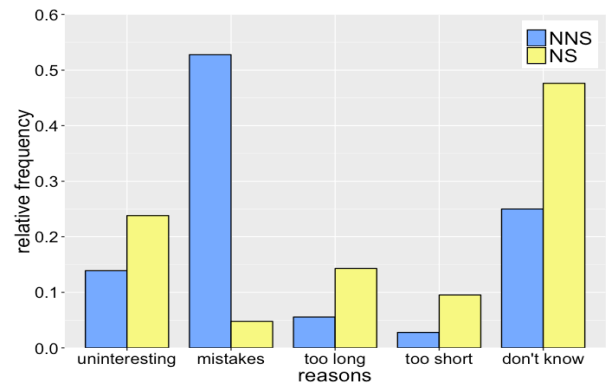

Figure 8: *Reasons given for e2 → e3 change in expression.*

H3: Change in complexity

NNS significantly changed the complexity of their sentences after s3 (*M*=-1.46, *SD*=2.85) in comparison to NS (*M*=.15, *SD*=2.45), *t*(45.94)=2.24, *p*=.03. NNS showed a propensity to reduce sentence complexity in s4, while NS did not, as shown in Figure 9.
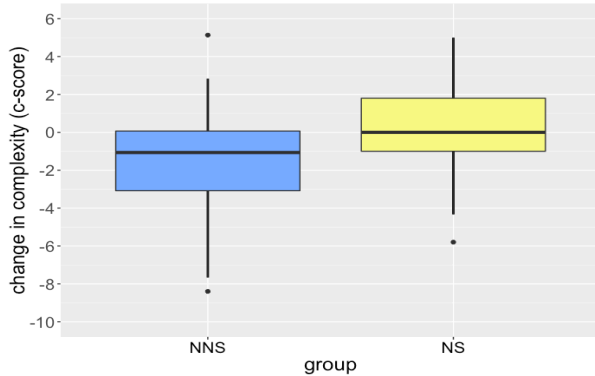

Figure 9: *Boxplot of change in sentence complexity (c-score) from s3 to s4.*

H4: Change in complexity by sensitivity

No difference was found between 'sensitive' (*N*=11, *M*=-1.77, *SD*=2.87) and 'not sensitive' (*N*=25, *M*=-1.33, *SD*=2.89) in their change in sentence complexity after s3, *t*(19.3)=-0.42, *p*=.7. This is shown in figure 10.
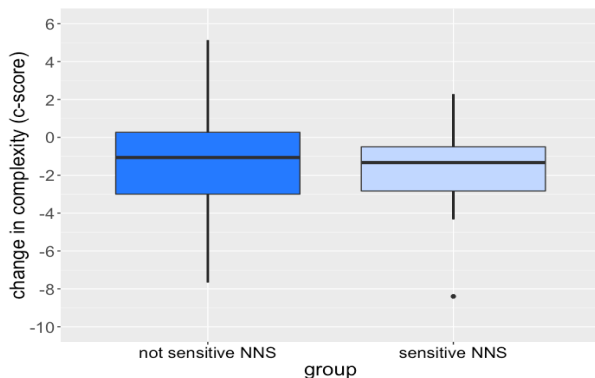

Figure 10: *Boxplot of change in learner s3-s4 sentence complexity by sensitivity to s4-s6 expression changes*

5. Discussion

The results of the study indicate that there is a difference in how NS and NNS interpret and respond to an e2 → e3 change in facial expression of an animated avatar as feedback to their typed sentences. NNS tend to view the change as resulting from their mistakes in vocabulary or grammar of their previous sentence, while this is absent in NS. The effect of this difference in interpretation is manifest in s4, where NNS reduce the sentence complexity, with the sample of NS showing no such reduction. No statistically significant difference was observed between the NS and NNS groups in the emotion attributed to this change in expression. Furthermore, there was no change in complexity of 'sensitive' vs. 'not sensitive' NNS.

There are several limitations of the experiment that could be addressed in future studies. The sample of NNS included a mix of nationalities and L1. There may be cultural differences and varying CLI effects between these groups in their response to facial expressions [33]. Studying independent populations who share nationality and L1 would allow us to identify any such differences. Three participants selected 'happy' to describe the e2 → e3 change in expression, which possibly indicates a misunderstanding, a lack of concentration, or a browser issue at the time of participation (despite prior testing). As participants undertook the experiment online at a time and place of their choosing, we cannot be certain as to the conditions under which the tasks were completed. Inviting participants into the lab and giving instructions in person may increase concentration and understanding in this regard, but it would impact privacy and anonymity.

A dissatisfaction with the limitations of the screen in creating 'deeper' interactions has, in part, spurred research into using robots in CALL. However, some of the robots presently used in language learning are unable to display changes in facial expression (e.g. Nao [7]). This paper makes a valuable contribution to the field of emerging language learning technologies by demonstrating that there is a great deal of potential in using expressions as a mechanism to provide feedback with currently available screen technologies. This is also a feature which could be incorporated into developing robotic systems (e.g. Furhat [34]) in the CALL field.

6. Conclusions and Future Work

The results presented in the previous section offer intriguing possibilities for the potential of using an avatar's change in facial expressions to guide language learners toward advantageous learning strategies. The e2 → e3 expression could be used to indicate to learners that the sentence they have just produced is not well-formed. The facial expression does not explicitly inform the user that what they produced is incorrect in the same way a red cross or harsh klaxon sound does, it only indicates that they are causing the listener some difficulty in understanding. It is then up to the user to respond, and as evidenced in this study, the reaction tends towards a reduction in sentence complexity. The extent to which this is manifest in the L2 production of learners with different abilities, cultural backgrounds and L1 is an avenue of research which has the potential to provide fine-grained insights into the language learning process. The growing corpus of learner sentences, errors and corrections will also allow for a gradual increase in automation of corrections. What will initially be a manual process involving a NS quickly assessing sentences and selecting an appropriate emotional response or correction, will develop into a semi-automated process where NS judgement and correction is needed only for sentences which contain structures/ambiguities not easily categorised by the corpus-based language model. On the learner side, by supplementing the e2 → e3 'error' expression with positive changes for well-formed sentences, intermediate language learners may be guided by facial expressions towards advanced levels and more confident production. This is the premise which underpins the novel, personalised e-learning platform currently under development at UCD.

7. Acknowledgements

8. References

[1] J. M. G. Alvarez and A. M. Ora, *MeryProject.com*, (2016) [online]. Available: http://www.meryproject.com.

[2] R. Lyster, S. Kazuya, and S. Masatoshi, "Oral corrective feedback in second language classrooms," *Language Teaching,* vol. 46, no. 1, pp.1-40, 2013.

[3] S. Carroll and M. Swain, "Explicit and implicit negative feedback," *Studies in Second Language Acquisition,* vol. 15, no. 3, pp. 357-386, 1993.

[4] A. Mackey, "Input, interaction and second language development: an empirical study of question formation in ESL," *Studies in Second Language Acquisition,* vol. 21, no. 4, pp. 557-587, 1999.

[5] E. M. Golonka, A. R. Bowles, V. M. Frank, D. L. Richardson, and S. Freynik, "Technologies for foreign language learning: a review of technology types and their effectiveness," *Computer Assisted Language Learning,* vol. 27, no. 1, pp. 70-105, 2014.

[6] S. Yun, J. Shin, D. Kim, C. G. Kim, M. Kim, and M. Choi, "Engkey: Tele-education robot," in *Third International Conference on Social Robotics,* 2011, pp. 142-152.

[7] T. Belpaeme, J. Kennedy, P. Baxter, P. Vogt, E. E. J. Krahmer, S. Kopp, K. Bergmann et al, "L2TOR-Second language tutoring using social robots," in *Proceedings of the ICSR 2015 WONDER Workshop, Paris, France,* 2015.

[8] D. Leyzberg, S. Spaulding, M. Toneva, and B. Scassellati, "The Physical Presence of a Robot Tutor Increases Cognitive Learning Gains," in *Proceedings of the Cognitive Science Society 34,* 2012, pp. 1882-1887.

[9] J. Kennedy, P. Baxter, and T. Belpaeme, "The robot who tried too hard: Social behaviour of a robot tutor can negatively affect child learning," in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction,* 2015, pp. 67-74.

[10] A. J. C. Sharkey, "Should we welcome robot teachers?" *Ethics and Information Technology,* vol. 18, no. 4, pp. 283-297, 2016.

[11] Pew Research Center, "Smartphone Ownership and Internet Usage Continues to Climb in Emerging Economies," 2016. [online] Available: http://www.pewglobal.org/2016/02/22/smartphone-ownership-and-internet-usage-continues-to-climb-in-emerging-economies/

[12] L. Jolin, "Fom Busuu to Babbel, language-learning startups adapt to thrive," 2017. [online] Available: https://www.theguardian.com/small-business-network/2017/mar/07/busuu-babble-language-learning-startups-adapt-thrive

[13] R. Ellis, "Input, Interaction and Second Language Acquisition," in *Understanding Second Language Acquisition*, 2nd ed, Oxford, Oxford university press, 2015.

[14] K. M. Bailey, "Competitiveness and anxiety in adult second language learning: Looking at and through the diary studies," in *Classroom Oriented Research in Second Language Acquisition*, H.W. Seliger and M. H. Long, Newbury House Publishers, Rowley, MA, 1983, pp. 67-103.

[15] P. D. MacIntyre and R. C. Gardner, "Language anxiety: Its relationship to other anxieties and to processing in native and second languages," *Language Learning,* vol. 41, no. 4, pp. 513-534, 1991.

[16] J. A. Russell, "Is there universal recognition of emotion from facial expressions? A review of the cross-cultural studies," *Psychological Bulletin,* vol. 115, no. 1, pp. 102-141, 1994.

[17] M. C. Frank, E. Vul, and S. P. Johnson, "Development of infants' attention to faces during the first year," *Cognition,* vol. 110, no. 2, pp. 160-170, 2009.

[18] E. Heyselaar, P. Hagoort, and K. Segaert. "In dialogue with an avatar, language behavior is identical to dialogue with a human partner," *Behavior research methods,* vol. 49, no. 1, pp. 46-60, 2017.

[19] K. Bergmann, H. P. Branigan, and S. Kopp. "Exploring the alignment space–lexical and gestural alignment with real and virtual humans," *Frontiers in ICT,* vol. 2, 2015.

[20] J. Hyde, E. J. Carter, S. Kiesler, and J. K. Hodgins, "Using an interactive avatar's facial expressiveness to increase persuasiveness and socialness," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems,* 2015, pp. 1719-1728.

[21] M. P. Stevenson and M. Liu, "Learning a language with Web 2.0: Exploring the use of social networking features of foreign language learning websites," *CALICO Journal,* vol. 27, no. 2, pp. 233-259, 2010.

[22] T. Geller, "Overcoming the uncanny valley," *IEEE computer graphics and applications,* vol. 28, no. 4, pp. 11-17, 2008.

[23] E. Zell, C. Aliaga, A. Jarabo, K. Zibrek, D. Gutierrez, R. McDonnell, and M. Botsch, "To stylize or not to stylize?: the effect of shape and material stylization on the perception of computer-generated faces," *ACM Transactions on Graphics (TOG),* vol. 34, no. 6, pp. 184-197, 2015.

[24] A. M. Glenberg, J. L. Schroeder, and D. A. Robertson, "Averting the gaze disengages the environment and facilitates remembering," *Memory & Cognition,* vol. 26, no. 4, pp. 651-658, 1998.

[25] G. Doherty-Sneddon, and F. G. Phelps, "Gaze aversion: A response to cognitive or social difficulty?" *Memory & Cognition,* vol. 33, no. 4, pp. 727-733, 2005.

[26] L. L. Lee and S. M. Canter, "Developmental sentence scoring: A clinical procedure for estimating syntactic development in children's spontaneous speech," *Journal of Speech and Hearing Disorders,* vol. 36, no. 3, pp. 315-340, 1971.

[27] H. S. Scarborough, "Index of productive syntax," *Applied Psycholinguistics,* vol. 11, no. 1, pp. 1-22, 1990.

[28] D. Larsen-Freeman, "An ESL index of development," *TESOL Quarterly*, vol. 12, pp. 439-448, 1978.

[29] L. Ortega, "Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing," *Applied Linguistics,* Vol. 24, no. 4, pp. 492-518, 2003.

[30] X. Lu, "Automatic analysis of syntactic complexity in second language writing," *International Journal of Corpus Linguistics,* vol. 15, no. 4, pp. 474-496, 2010.

[31] K. W. Hunt, "Grammatical Structures Written at Three Grade Levels," NCTE Research Report, No. 3, 1965.

[32] R. A. Fisher, "The logic of inductive inference," *Journal of the Royal Statistical Society,* vol. 98, no. 1, pp. 39-82, 1935.

[33] R. E. Jack, O. G. B Garrod, H. Yu, R. Caldara, and P. G. Schyns, "Facial expressions of emotion are not culturally universal," *Proceedings of the National Academy of Sciences,* vol. 109, no. 19, pp. 7241-7244, 2012.

[34] S. A. Moubayed, G. Skantze, and J. Beskow. "The furhat back-projected humanoid head–lip reading, gaze and multi-party interaction," *International Journal of Humanoid Robotics,* vol.10, no. 1, 2013.