Can a robot companion help students learn Chinese tones? The role of speech and gesture cues

Zinina Anna A.[1,2,3] [0000-0001-9575-1875], Kotov Artemiy A.[1,2] [0000-0003-3353-5549],

Arinkin Nikita A.[1,2] [0000-0003-2303-2817], Gureyeva Anna N. [3]

[1] National Research Center "Kurchatov Institute", Moscow, Russia
[2] Russian State University for the Humanities, Moscow, Russia
[3] Moscow State Linguistic University, Moscow, Russia
zinina_aa@nrcki.com

Abstract. The paper examines the possibility of using companion robots (F-2 robot, as the example) to assist humans in learning the tonal system of Chinese. The experiment compares the effectiveness of supportive cues via speech and gestures. The results of the experiment (N=20, 4 males, mean age 24,5) show that the learning is effective – subjects learn the pronunciation of Chinese syllables with the given tones. At the same time the effectiveness of speech cues is higher than of the gestural cues. It is the speech cue that seems to be more understandable during the subjects pass the introduction to the system of Chinese tones. In general, a combination of conditions in the experiment is effective when subjects learn Chinese phonetics with gestural cues and with speech cues – in any order of conditions. In addition, the robot with gesture cues is perceived by the participants as more natural and engaging, eliciting more sympathy. The results suggest that companion robots with gestural behavior can be used to support the educational process and increase student engagement.

Keywords: Social Robots, Learning, Human-Machine Interaction, Chinese

1 Introduction

The phonetics of the Chinese language is one of the most difficult aspects in its study, since it is here that a large number of linguistic phenomena that have no analogues in Indo–European languages are concentrated. Tone in Chinese has a semantic function: a syllable consisting of a given set of phonemes, when pronounced in different tones, corresponds to different characters and has different meanings. Different dialects of Chinese have diverse tonal systems. In normalized Mandarin *Putonghua* there are four tones, contrasted by pitch and voice movement pattern: (a) high, even; (b) rising; (c) falling-rising; (d) falling. For example, the syllable *tu* has at least four different lexical meanings depending on its tonal pattern: the first tone tū would mean 'suddenly' (突), the second tú would mean 'image' (图), the third tǔ would mean 'soil' (土), and the fourth tù would mean 'hare' (兔). The fifth, or zero tone is not considered as independent, as it is pronounced neutrally and is not opposed to other tones neither phonetically nor semantically [Panov, 1967; Speshnev, 1980; Susov, 2008]. A description of the voice patterns in Chinese tones can be represented in Fig. 1.

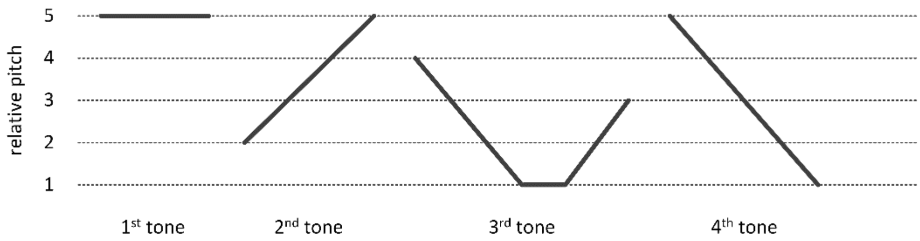


Fig. 1. Voice frequency patterns of Chinese tones.

The melody of the first tone is high, smooth, long, with regular intensity and only some decrease to the end (it may give an impression of an unfinished statement to a European listener). The melody of the second tone is short, rapidly rising, with maximum intensity at the end of the syllable; it gives an impression of interrogation. The third tone is low, long, and has a falling-rising pattern, with maximum intensity on the low note. The fourth tone is short, sharply falling from the highest point to the lowest. The falling tone is accompanied by a sharp decrease of intensity (the melody of the fourth tone gives an impression of a categorical command) [Zadoenko, Huang, 1993; Aleksakhin, 2006].

The phonetic system of Chinese differs significantly from European languages. Speakers of tonal and non-tonal languages have significantly different abilities to learn the phonetics of Chinese [Francis et al., 2008]. Therefore, when European students learn Chinese, phonetics is given significant time in the first months of the curriculum. As a rule, European students do not pay enough attention to the change of pitch: in most cases speakers of non-tone languages use pitch only for lexical and phrase emphasis, to express emotions or to choose the focus of a statement [Malyshev, Kiselevich et al., 2018]. In the process of learning, students should pay attention to the tone contour and learn to recognize and pronounce Chinese syllables with a given tone. To do this, the teacher must use special cues to direct the student's attention to phonetic processes. McGuiness [1997] emphasizes the difficulties of learning tones in Chinese, noting that most native English learners of Chinese consider the concept of using tones to change word meaning as so distant from their native English that learning tones becomes problematic, neglected, or both. In [Xu, 2011] was shown that the presence of tones is the linguistic aspect that causes the most difficulty for learners – for example, when pronouncing tones, students confuse them with stresses. It is also noted that a common mistake is the incorrect pitch of the beginning of the tones and the addition of extra stress to Chinese words, leading to a disruption of syllable equivalence. In the survey on the evaluation of difficulties in learning Chinese [Chiang, 2002], the students who participated in the survey – 26 second-year students – most often mentioned difficulties with the study of tones: students would like to get more explanations on this topic at the elementary level, in the first year of study.

1.1 The role of gestures in learning a foreign language

To solve the difficulties in learning Chinese phonetic system, teachers can use different exercises and means to direct students' attention to phonetic phenomena, for example – use iconic gestures [Francis et al., 2008]. Gestures can enhance comprehension and memorization by supporting verbal messages, helping to illustrate concepts, conveying additional information, and providing cues. Indirect benefits of using gestures are capturing the listener's attention and creating a comfortable, trusted atmosphere. Therefore, teachers often use gestures in the classroom, even when explaining abstract concepts or phonetic phenomena that have no spatial representation. This directs the attention of the learners. The interaction of words and gestures is important, as well as their optimal ratio: a lack of gestures by the lecturer gives the impression of stiffness and insecurity. Their excessive use is perceived as unnatural and intrusive behavior.

Iconic gestures are a particular type of the iconic signs, where the signifier is in some aspect similar to the signified [Peirce, 2000]. Iconic gestures can denote an action, an event, or an object by reproducing the way of performing the action, the trajectory of the action, the shape of the object, or the way of interacting with the object (affordance). Iconic gestures are natural and common part of spoken language, visually representing the concepts to which they refer [Peirce, 2000]. We can consider Chinese tone gestures as iconic: in such a gesture, a person may represent the shape of the tone outlined with his hand, where the height at which the hand is positioned denotes the pitch of the sound. In addition to conveying iconic meaning, gestures can also be used to represent abstract metaphors [McNeill, 1992], and in these circumstances the gesture is classified as metaphorical. For example, if a student says that his grades have improved and at the same time points upward, he is using a metaphorical gesture where the upward movement symbolizes the improvement. From this perspective, the iconic gesture indicating tone can be seen as metaphorical because of the abstract correspondence between the height of the hand placement and the height of the sound.

Iconic gestures are known to affect the learning of foreign words [Kelly et al., 2009; Hirata et al., 2014; Dargue et al., 2019]. There is also evidence that people use head and hand gestures to project the tones of hieroglyphs, which makes it easier to distinguish similar-sounding hieroglyphs [Chen, Massaro, 2008; Baills et al., 2019; Morett, Chang, 2015].

1.2 Using robots for foreign language learning

In this study, we used a companion robot F-2 that interactively assisted humans in learning the tonal system. It is known that robots have long been used in education [Golonka et al., 2014; Young et al., 2012; van Ewijk, 2020]. Physically embodied robots can be used for educational programs that require interaction with the surrounding physical world; they are able to add social interaction to learning, that stimulates student engagement in the educational process, increasing its effectiveness [Kidd, Breazeal, 2004; Han et al., 2008]. Physically embodied robots are perceived with more enthusiasm and interest, as compared to virtual agents. Also, with them, users show higher learning results [Belpaeme et al., 2018]. With their ability to support natural communication, robots are an effective tool for language learning [van den Berghe et al., 2019; Belpaeme et al., 2018; Kennedy et al., 2016; Alemi et al., 2014]. This rapidly growing area is referred to as – *Robot-Assisted Language Learning* (RALL) [van den Berghe et al., 2019]. Although this technological field is quite young, it is also very promising: social robots seem extremely appropriate for use in language learning due to their inherently social design [Randall, 2019].

There is not much experience in the use of robots for learning Chinese [Hao, 2012]. There are studies indicating the influence of iconic [de Wit et al., 2018] and rhythmic [Gluhareva, Prieto, 2017] robot gestures on foreign word memorization. However, there is practically no research on whether robots as teachers using head and hand gestures (i.e., pitch indicators) can improve the learning of different tones in Chinese. In [Zhang, de Haas, 2020], there are some preliminary results indicating the effectiveness of gesture cues in learning Chinese tones. The paper showed that the robot, which does not use gestures when teaching Chinese tones, increased the efficiency of learning, but only if the participants were previously introduced to the Chinese characters by the robot using gestures. In general, the authors do not deny the prospects of using a social robot to learn Chinese tones. In our study, we will not only consider the impact of the social robot on learning Chinese tones, but also compare the iconic gesture-based cueing of Chinese tones with the voice-based cueing.

2 Experimental research

As part of the study, an experiment was conducted in which students were learning Chinese tones using two F-2 robots. Both robots explained the concept of Chinese tones, and the explanation of one of them (condition 1) was supported by gestures illustrating the pattern of voice movement when playing the tone, while the other explained it only in speech (condition 2). For the following questioning, the robots were differentiated by labeling: the robot with gesture cues was marked with a triangle, and the robot using only speech cues was marked with a square.

Each subject learned tones with each of the robots – in random order. Each robot's learning session consisted of two stages – the introduction stage and the training stage. In the introduction stage, the robot briefly explained the concept and function of tones in Chinese, and then – demonstrated the difference of tones by reproducing one syllable in different tones. The robot's speech was synthesized using Yandex Speech API (state of the art text synthesis software), and the syllables in Chinese were recorded by a native speaker – university teacher. Both robots used speech description of tone (*Here is the first tone. The voice is high, goes straight...*) when demonstrating syllable differences. A robot using gestures also demonstrated a gesture depicting the tone pattern. Also, both robots voiced the translation of each tonal syllable, thereby illustrating the semantic differentiating nature of the tones. Then, at the training stage, the students were asked by the robot to pronounce given syllables themselves with a specific tone. Further, Fig. 2 shows the structure of the tone training stage.

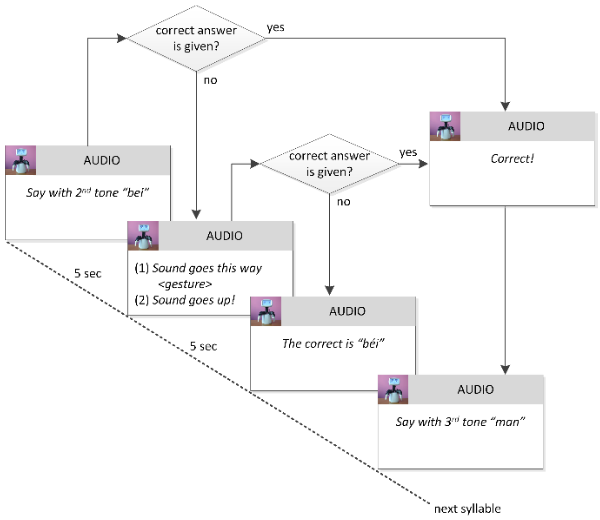


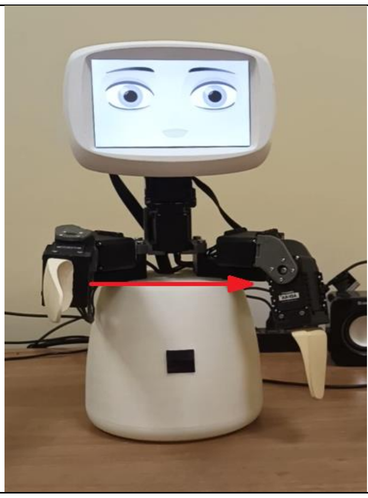
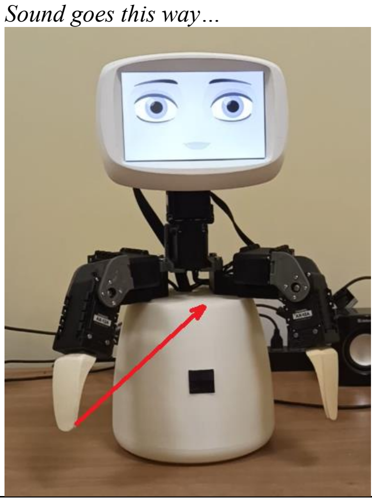
Fig. 2. Schematic representation of the training stage

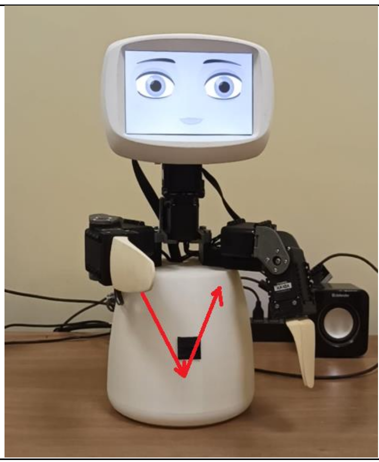
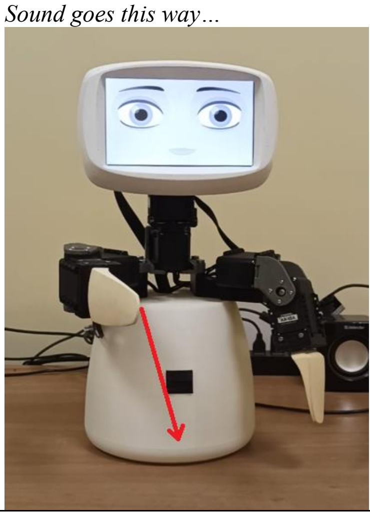
The scheme shows that the robot sequentially named different syllables and asked the subject to pronounce them with a certain tone (*Say the first tone "ma"...*). When an incorrect answer was received, or in case the answer was not received within 5 seconds, the robot played a hint: the robot with gestures used a gesture to represent the voice movement of the required tone as a cue (condition 1). The other robot (condition 2) illustrated the required voice movement in speech, for example: *The voice is high, it goes straight*, *The voice falls*, etc. If the participant then made a mistake in tone or did not give an answer within the next 5 seconds, an audio recording with the correct pronunciation by the native speaker was played. Table 1 shows the robot's gestures and actual cues for a particular tone of Chinese.

Table 1. Robot cues in the training phase for two experimental conditions.

	Condition 1. A robot using gestures as a cue[1]	Condition 2. A robot using speech as a cue
The first tone:	*Sound goes this way…*	*The voice is high, and goes straight.*

[1] Robot with the square is represented here for photos only. The actual performance in this condition was by the robot with triangle.

		
The second tone:	*Sound goes this way…* 	*The voice rises from the bottom up*
The third tone:	*Sound goes this way…*	*The voice falls smoothly down, then up*

		
The fourth tone:	*Sound goes this way…* 	*The voice falls straight down*

The syllable sets used by the robots in the explanation and training stages differed between the two robots, but were similar in structure and complexity (example: *lan* – *man*, *mu* – *lu*; with different initials but similar or identical finals and identical tones). The syllable sets had the same size, 20 units. In the training stage, the student first had to pronounce the syllable suggested by the robot alternately in four tones. First, a syllable from the training stage was used (in the first condition, the syllable *ma*; in the second condition, the syllable *lu*), then another syllable (in the first condition, the syllables *du*; in the second condition, the syllables *mu*). After that, the robot instructed the student to pronounce the suggested syllable in a new tone (syllables and tones were given in different order, but the syllable structure and tone were similar for the two conditions). Correctness of pronunciation was controlled by an assistant experimenter (a specialist in Chinese linguistics) from another room according to the Wizard-of-Oz scheme. During the evaluation, special attention was paid to the tonal pattern of the syllable. Due to the fact that the pronunciation rules in Chinese and Russian differ significantly, during the experiment subjects were allowed to pronounce inaccurately the syllables suggested by the robot (i. e. with incorrect pronunciation of the initial – the starting consonant sound of a syllable, or of the final – the end of it). It was the correctness of the tonal pattern that was in the focus of attention.

A within-subject experimental design was used: each participant of the experiment was learning from both robots in turn, and the order in which the learning took place was randomized in order to achieve maximum objectivity in the evaluation. In the experimental instruction, the investigator explained to the subject that both robots were to help the person learn Chinese phonetics, and the subject would be trained with each robot to compare them afterwards. After completing the training with each robot, the subject filled out a questionnaire to rate their impressions using a semantic differential scales and chose the preferable robot.

The subjects were students – 20 people (men 4, women 16; mean age – 24,5 years). None of the subjects had studied Chinese before. Each experimental interaction was recorded on camera.

3 Results

On the basis of summarized statistics, we can conclude that the learning of Chinese tones is more effective with the robot using speech cues (Wilcoxon matched pairs test, $p < 0{,}005$). However, it is more correct to continue further analysis considering the order of presentation. It can be observed that the efficiency of correct tone reproduction increases by the second presentation, regardless of the type of condition – hence, the learning of Chinese tones was successful. The difference in the progress of correct answers is statistically significant (Wilcoxon matched pairs test, $p < 0{,}001$).

When analyzing the data, we can also observe that if the first condition was with a gesture cue, the learning efficiency was slightly lower than if the first condition was with a speech cue, this difference is significant (Wilcoxon matched pairs test, $p < 0{,}001$).

If we look at the graphical representation of the results (Fig. 3), we can conclude that the verbal cues work better: test subjects recognize and understand the voice cue quickly, while the gesture cue is sometimes misread. The graph shows a trend: in the first presentation, the gesture cue works only 8% of the time, while the speech cue works 20,5% of the time.

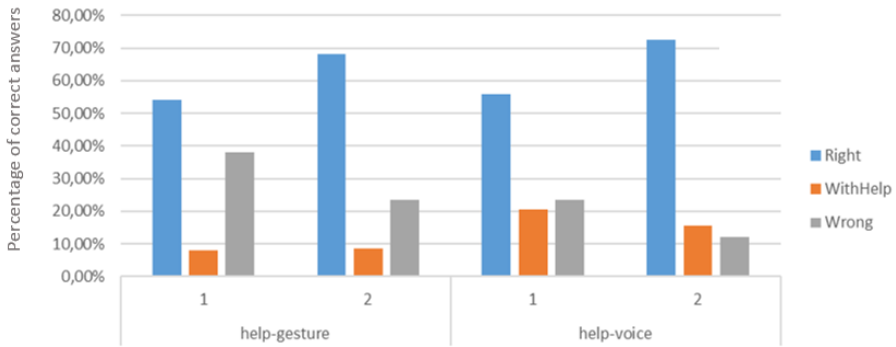


Fig. 3. Percentages of *Right*, *Wrong*, and *WithHelp* (after a hint) answers in the first (1) and second (2) presentations of the corresponding condition.

This can be compared to the fact that according to the post-experiment survey, 20% of participants perceived the iconic gestures depicting the tone contour as rhythmic gestures of the robot, which did not carry any semantic meaning, although the robot accompanied them with the words *Sound goes like this...* This implies the need to accompany gestural cues with more clear verbal explanations, focusing more attention on them.

The results of the repeated-measures ANOVA indicate that both the condition factor (gestures or speech) and the presentation factor (first and second presentation of the condition) are significant (Wilks lambda = 0,92146, $F(2, 397) = 16.920$, $p < 0,001$) for learning Chinese tones.

An analysis of the efficiency of learning each Chinese tone indicates that the subjects learned the 4th tone significantly easier, the worst of all the subjects learned the 1st tone (Fig. 5). In Chinese phonetics the 1st tone is considered the easiest to learn: it has an flat tonal pattern. However, according to the results of the experiment, we can say that the interaction with the robots with different types of cues is not enough for its learning. The supposed reason for this result is the insufficient separation by the robot of the phase of raising the arm to the starting point and the phase of the tone image (flat hand movement at a high level). Based on this, further research is required to attract more attention to the development of a gesture cue system for the robot.

Figure 4 also shows the effectiveness of different types of cues in learning different tones of Chinese: it can be concluded that the speech cue is significantly ($p < 0,05$) more effective for learning the 2nd tone. This may also be due to the fact that the speech cue *Sound rises up* is more effective than the diagonal raising of the robot hand, which students could perceive as an assistive movement and not as an iconic gesture.

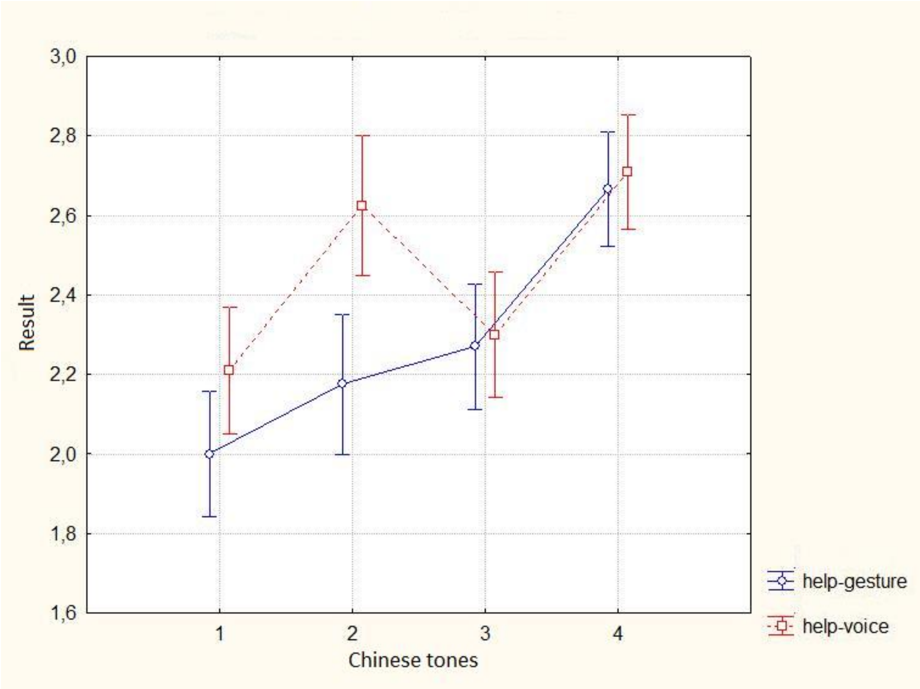


Fig. 4. The effectiveness of learning 1st, 2nd, 3rd and 4th Chinese tones depending on the cue type.

After the experiment, we asked the subjects if they noticed a difference between the robots and their behavior: only 10 subjects (50%) could correctly identify the difference between the two conditions. Some of the subjects perceived iconic gestures as rhythmic robot gestures that carried no semantic meaning. At the same time, a person's inability to describe the difference between robots in the aspect of gestural behavior does not mean that the gestures had no influence on them: robot gestures can influence a person, even if the person does not describe it in the final report [Zinina et al., 2019].

Based on the questionnaire data, it is not possible to identify the subjects' preferred type of cues: 11 subjects preferred learning with a robot that used verbal cues (55%), 9 people (45%) preferred a robot with gestural cues.

According to the questionnaire, the robot with gestural prompts seemed more engaging (2,95 out of 5 on the semantic differential scale) and comfortable (2,7 out of 5), but participants noted that this robot was also more distracting (1,9 out of 5 – compared to 1,6 out of 5 for the second condition). It was also found that subjects perceived its cues to be more confusing (1,6 points out of 5), as compared to the robot with speech cues (1,35 points out of 5).

The learning behavior of the robot with verbal cues was perceived by the subjects as more understandable (4,7 out of 5) and comfortable (4,35 out of 5) – the difference in this parameter was statistically significant (T-test, $p < 0{,}05$), see Fig. 5.

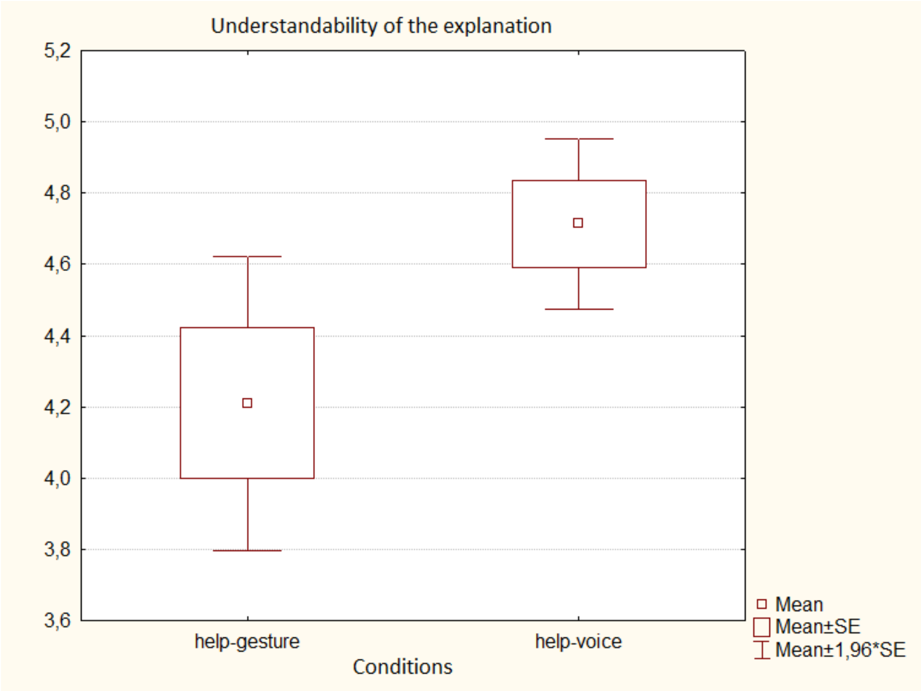


Fig. 5. Differences in the subjective evaluation of the understandability of the explanation of Chinese tones in the robot with gesture and speech cues.

According to the subjective evaluation of the effectiveness of the cues of each robot, the most effective were the speech cues (2,4 points out of 5).

A more detailed analysis of the results shows that the type of cues significantly affects the success specifically in the first introduction to the Chinese tones – this is confirmed not only in the evaluation of learning effectiveness, but also in the subjective evaluation by the subjects after interaction with the robots (Fig. 6) – (T-test, $p < 0,05$).

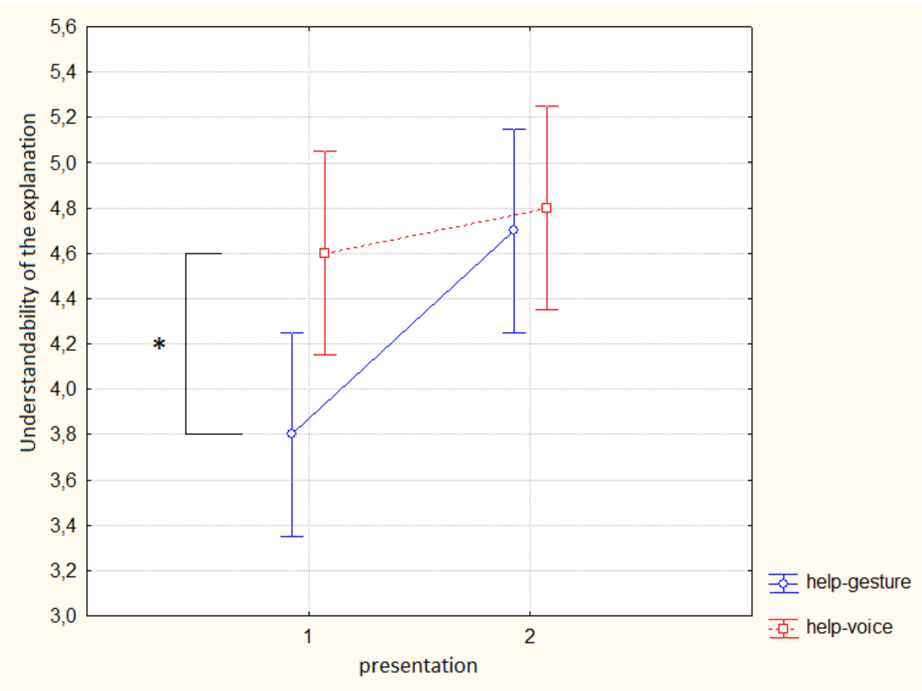


Fig. 6. Differences in subjective assessment of the understandability of the explanation of Chinese tones (a) in the robots with gestural or speech cues, (b) within the first and second presentation.

Based on their subjective evaluation of the understandability of the explanation of Chinese tones, subjects identified speech cues as more comprehensible – but the difference in the second turn of the presented conditions is small: it shows that subjects equally well learned to pronounce syllables in the correct tone, regardless of the order of the experimental conditions.

The insufficient effectiveness of gestural iconic cues in the first stage can be explained by the novelty effect when meeting the robot, the limitations of the robot's body model, the requirements to better separate the preparatory and iconic gesture stages, and the need to better focus the student's attention on the robot's iconic gesture. Based on the frequent misreading of gesture cues by test subjects, more attention should be paid at the design the robot gesture cue system, making it more explicit to the learners.

4 Conclusion

Based on the results of the experimental study, it has been demonstrated that the gestural iconic cues of companion robots are not as effective as the speech cues in the context of primary Chinese language teaching tones. However, the use of gestures allows the robot to better establish emotional contact with the learner: the robot with gestural cues was perceived by the participants as more natural and engaging, eliciting the greatest sympathy.

After passing the two conditions, the subjects learned the topic of Chinese tones equally well, and the learning efficiency did not depend on the order in which the conditions were presented. This may indicate that the use of verbal and gestural cues creates a complex effect that allows the student to learn the topic after passing the training with these two types of learning methods.

The most important result of the experiment was the confirmation of the fact that a social robot can be used to teach the Chinese tones. Further research will focus on finding the optimal learning strategies that specifically combine the different types of cues: speech and gestures. The result seems to be extremely important because learning Chinese is a complex process in which any means of optimization will be in demand. That is why it is important to develop training communication strategies for companion robots in this way, in order to maximize the effectiveness of learning with robots.

Acknowledgements

The project is in part supported by the Russian Science Foundation, project No 19-18-00547, https://rscf.ru/project/19-18-00547/

References

1. Aleksakhin, A.N. (2006) Theoretical phonetics of the Chinese language. Moscow: AST: Vostok-Zapad.
2. Alemi, M., Meghdari, A., Ghazisaedy, M. (2014). Employing humanoid robots for teaching English language in Iranian junior high-schools. International Journal of Humanoid Robotics, 11, 1450022-1–1450022-25. doi:10.1142/S0219843614500224
3. Baills, F., Suárez-González, N., González-Fuente S., Prieto, P. (2019) Observing and producing pitch gestures facilitates the learning of Mandarin Chinese tones and words. Studies in Second Language Acquisition, 41(1), 33-58.
4. Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. Science robotics, 3(21), eaat5954.
5. Chen, T. H., & Massaro, D. W. (2008). Seeing pitch: Visual information for lexical tones of Mandarin-Chinese. The Journal of the Acoustical Society of America, 123(4), 2356-2366.
6. Chiang, M. H. (2002). An investigation of students' perspective on Chinese language learning. Journal-Chinese Language Teachers Association, 37(1), 43-64.
7. Dargue, N., Sweller, N., & Jones, M. P. (2019). When our hands help us understand: A meta-analysis into the effects of gesture on comprehension. Psychological Bulletin, 145(8), 765.
8. de Wit, J., Schodde, T., Willemsen, B., Bergmann, K., De Haas, M., Kopp, S., ... & Vogt, P. (2018, February). The effect of a robot's gestures and adaptive tutoring on children's acquisition of second language vocabularies. In Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction (pp. 50-58).
9. Francis, A. L., Ciocca, V., Ma, L., & Fenn, K. (2008). Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. Journal of Phonetics, 36(2), 268-294.

10. Gluhareva, D., & Prieto, P. (2017). Training with rhythmic beat gestures benefits L2 pronunciation in discourse-demanding situations. Language Teaching Research, 21(5), 609-631.
11. Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. F. (2014). Technologies for foreign language learning: A review of technology types and their effectiveness. Computer Assisted Language Learning, 27 (1), 70–105. doi:10.1080/09588221.2012.700315
12. Han, J. H., Jo, M. H., Jones, V., & Jo, J. H. (2008). Comparative study on the educational use of home robots for children. Journal of Information Processing Systems, 4(4), 159-168.
13. Hao, Y. C. (2012). Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. Journal of phonetics, 40(2), 269-279.
14. Hirata, Y., Kelly, S. D., Huang, J., & Manansala, M. (2014). Effects of hand gestures on auditory learning of second-language vowel length contrasts. Journal of Speech, Language, and Hearing Research, 57(6), 2090-2101.
15. Kelly, S. D., McDevitt, T., & Esch, M. (2009). Brief training with co-speech gesture lends a hand to word learning in a foreign language. Language and cognitive processes, 24(2), 313-334.
16. Kennedy, J., Baxter, P., Senft, E., & Belpaeme, T. (2016, March). Social robot tutoring for child second language learning. In 2016 11th ACM/IEEE international conference on human-robot interaction (HRI) (pp. 231-238). IEEE.
17. Kidd, C. D., & Breazeal, C. (2004, September). Effect of a robot on user perceptions. In 2004 IEEE/RSJ international conference on intelligent robots and systems (IROS)(IEEE Cat. No. 04CH37566) (Vol. 4, pp. 3559-3564). IEEE.
18. Malyshev, G.I., Kiselevich, Y.E., Mitchell, P.D. (2018) Difficulties in studying phonetics of Chinese by Russian-speaking students: the main mistakes and the ways to fix them. Bulletin of the Tambov University. Humanities Series, 23(173), 43-48.
19. McGinnis, S. (1997). Tonal spelling versus diacritics for teaching pronunciation of Mandarin Chinese. The Modern Language Journal, 81(2), 228-236.
20. McNeill, D. (1992) Hand and mind: What gestures reveal about thought. Chicago, IL: The University of Chicago Press.
21. Morett, L. M., & Chang, L. Y. (2015). Emphasising sound and meaning: Pitch gestures enhance Mandarin lexical tone acquisition. Language, Cognition and Neuroscience, 30(3), 347-353.
22. Panov, M.V. (1967) Russian phonetics. M.: Prosveshcheniye.
23. Peirce, Ch. (2000) The Beginnings of Pragmatism. Translation from English, foreword by V. V. Kiryushchenko, M. V. Kolopotin. St. Petersburg: Laboratory of Metaphysical Research of the Philosophical Faculty of St. Petersburg State University; Aletheia.
24. Randall, N. (2019). A survey of robot-assisted language learning (RALL). ACM Transactions on Human-Robot Interaction (THRI), 9(1), 1-36.
25. Speshnev, N.A. (1980) Phonetics of the Chinese language. L.: Leningrad University Publisher.
26. Susov, I. P. (2008) Introduction to Linguistics: Textbook for Students of Linguistic and Philological Specialties. Tver State University. Moscow: AST: Vostok - Zapad.
27. Van den Berghe, R., Verhagen, J., Oudgenoeg-Paz, O., Van der Ven, S., & Leseman, P. (2019). Social robots for language learning: A review. Review of Educational Research, 89(2), 259-295.
28. van Ewijk, G., Smakman, M., & Konijn, E. A. (2020, June). Teachers' perspectives on social robots in education: an exploratory case study. In Proceedings of the interaction design and children conference (pp. 273-280).

29. Xu, S. Y. (2011) Problems of teaching phonetics of Chinese language to Russian students. Scientist, 12(35), 141-144.
30. Young, M. F., Slota, S., Cutter, A. B., Jalette, G., Mullin, G., Lai, B., & Yukhymenko, M. (2012). Our princess is in another castle: A review of trends in serious gaming for education. Review of educational research, 82(1), 61-89.
31. Zadoenko, T.P., Huang, Sh. (1993) Fundamentals of Chinese. An Introductory Course, 2nd edition. Moscow: Nauka, Oriental Literature Publishing House.
32. Zhang, P., & de Haas, M. (2020, March). Effects of pitch gestures on learning Chinese orthography with a social robot. In Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction (pp. 180-182).
33. Zinina, A., Arinkin, N., Zaydelman, L., Kotov, A. (2019) The Role of Oriented Gestures during Robot's Communication to a Human. Computational Linguistics and Intellectual Technologies, 18 (25), 800-808.