PICTALKY: Augmentative and Alternative Communication Software for Language Developmental Disabilities

Anonymous ACL submission

Abstract

Children with language disabilities face communication difficulties in social life. They are often deprived of the opportunity to participate in social activities due to their difficulty in understanding or using natural language. In this regard, Augmentative and Alternative Communication (AAC) is a practical means of communication for children with language disabilities. In this study, we propose PICTALKY, which is an AI-based AAC system that helps children with language developmental disabilities to improve their communication skills and language comprehension abilities. PICTALKY can process both text and pictograms more accurately by connecting a series of neural-based NLP modules. Moreover, we perform quantitative and qualitative analyses on the essential features of PICTALKY. It is expected that those suffering from language problems will be able to express their intentions or desires more easily and improve their quality of life by using this service. We have made the models freely available alongside a demonstration of the Web interface [1]. Furthermore, we implemented robotics AAC for the first time by applying PICTALKY to the NAO robot.

1 Introduction

The majority of people with language disabilities suffer in their daily lives as they cannot understand or speak the language, which is a means of communication. Therefore, they may be deprived of the opportunity to participate in social activities and they may experience financial difficulties. In general, people with speech disorders have lower employment rates than people with other types of disabilities. Moreover, the proportion of people with autism disorder has been increasing every year (Zablotsky et al., 2019), and accordingly, a solution is required.

[1] https://pictalky-311.de.r.appspot.com/

Augmentative and alternative communication (AAC) has been suggested and applied to solve communication problems for people with language disabilities (Beukelman et al., 1998). This approach enables nonverbal communication instead of language. Although several AAC software resources are available, existing software packages are expensive, difficult to use, and only provide simple functions. To address these problems, we present a novel AAC system for children with language developmental disabilities. We refer to our AAC software as PICTALKY. Neural-based grammar error correction (GEC) and a symbol-based text-to-pictogram (TP) module are utilized in our model. Thus, PICTALKY offers neural- and symbol-based AAC for the improvement of communication and language learning, which have not been used in existing software.

From the perspective of NLP, the speech errors from people with language disabilities can be interpreted as grammatical errors at the morphological and syntactic levels. To handle these errors, neural GEC is loaded into PICTALKY. Moreover, we consider both text and image processing for AAC education and communication. After a sentence is entered as input through the speech-to-text (STT) module, it is passed through the neural GEC and natural language understanding (NLU) modules. Finally, the corresponding pictograms are displayed.

Our proposed service is aimed at children aged 0 to 14 years who have language developmental disabilities caused by intellectual or autism disabilities. The first reason that we focus on children is that early treatment during childhood is critical. According to Lenneberg (1967), language must be acquired during a critical period that ends at approximately the age of puberty with the establishment of the cerebral lateralization of function. Unless language is learned during this period, it is difficult for language to be used freely. This may result in social deterioration, contraction, aggression, and other problematic behaviors, which eventually affect the overall quality of life and satisfaction of the person (Schwarz et al., 2001).

The second reason is that there is currently insufficient social support for language therapy. Not all children with developmental disabilities can benefit from public systems owing to the limited support. Moreover, in addition to the children, their family and caregivers them experience difficulties.

Therefore, we propose PICTALKY, which complements the limitation of existing products and increases the accessibility of children with language disabilities to appropriate education and treatment. We expect that not only the people with language disabilities but also their caregivers can have more easier education and communication by using this service. Furthermore, in addition to the implementation in the Web application, we apply PICTALKY to the NAO robot, thereby providing the first robotics AAC. We expect that robotics AAC can draw interest of children, so that they can use AAC more friendly and easily.

Our contributions are as follows:

- We propose PICTALKY for people with language diabilities, which is the first AAC software with GEC and a synonym-replacement system for accurate language processing.
- We analyze each detailed function of PICTALKY quantitatively and qualitatively. Also, we measure the satisfaction score during the actual services.
- We present a novel metric known as text-to-pictogram accuracy (TPA) to measure the performance of converting texts into pictograms.
- We open PICTALKY in the form of a platform, so that it can help people with language disabilities and contribute to the research in this area.
- We implement robotics AAC for the first time by applying PICTALKY to the NAO robot.

2 Background

2.1 Language Developmental Disabilities

Language disorder is a slow-speech phenomenon due to late development of the speech center in the brain (Tomblin et al., 2003). Language disorders can be categorized into four main categories: expressive language disorder, mixed receptive-expressive language disorder, phonological disorder, and stuttering.

People with expressive language disorder have relatively normal receptive language ability to understand other people's words, but difficulty in language expression. They tend to replace simple words or sentences with gestures. People with mixed receptive-expressive language disorder shows a disability in understanding other people's words and in expressing their thoughts in language. In phonological disorder, there is a common occurrence of incorrect pronunciation in consonants, especially mispronouncing consonants or omitting the coda (auslaut) of syllables. Most frequently mispronounced consonants are [s], [z], [ʃ], [ʒ], etc, also, there are mispronunciations in vowels too. The speech of people with stuttering is cut off abnormally often or the speed of which is irregular. The repetition of sounds or syllables, extension of speech sounds, and blockage of speech can be observed. Also, their speech typically begins by repeating the first consonant of a phrase. Children generally do not recognize stuttering, as getting older, they become aware of their speaking problems, and emotional reactions occur to avoid being not fluent.

These disorders can be interpreted as grammatical errors at morphological and syntax levels from the perspective of natural language processing. Thus, deep learning-based grammar error corrector has been developed and loaded into PICTALKY's software .

2.2 AAC Software for Language Developmental Disabilities

Several AAC software platforms have been developed for language education. TouchChat[2] is a symbol- and text-based AAC tool with a text-to-speech (TTS) service. AVAZ[3] is a language education service that uses pictograms. TalkingBoogie (Shin et al., 2020) is software that supports the caregivers of children.

Systems that use AAC have also been developed for communication in daily life. Proloquo2Go[4] and QuickTalkAAC[5] enable people to communicate by

[2] https://touchchatapp.com/
[3] https://www.avazapp.com/
[4] https://www.assistiveware.com/products/proloquo2go/
[5] https://digitalscribbler.com/quick-talk-aac/

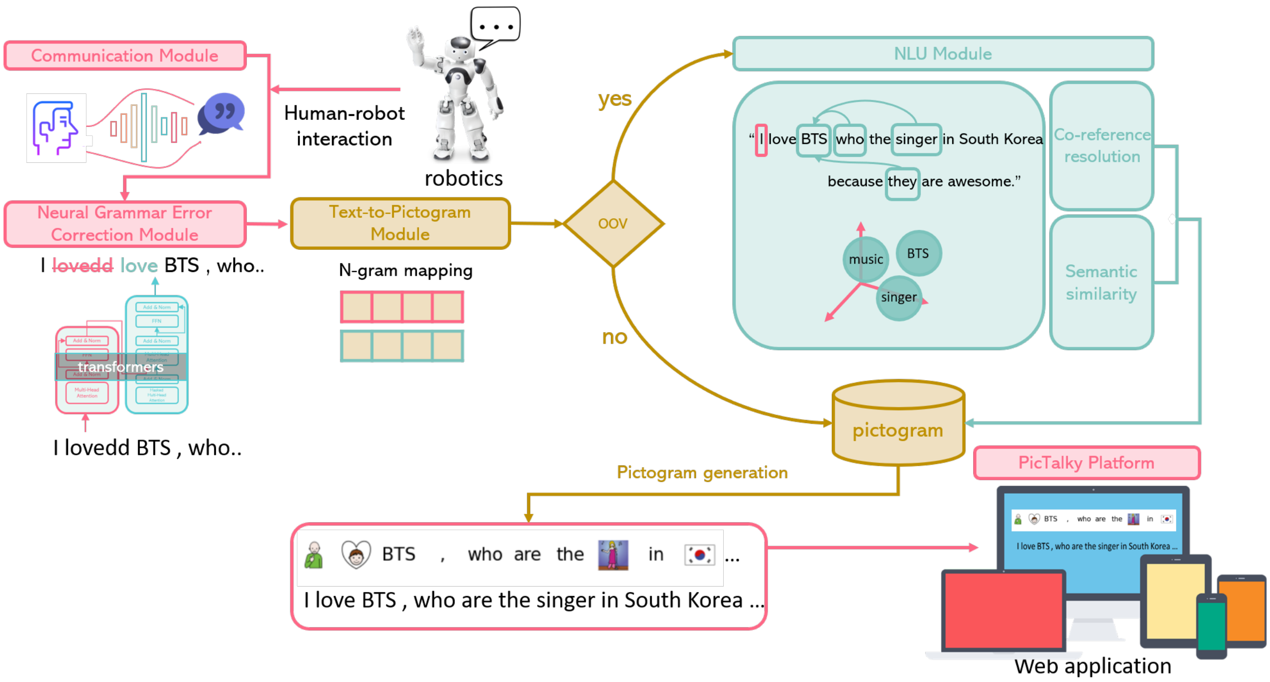


Figure 1: Overall architecture of PICTALKY.

using symbols or text with a TTS service. iCommunicate[6] is a visual and text AAC application that allows for the creation of pictures and storyboards. Although several AAC software platforms have been developed, certain problems remain, such as difficulty of use and high costs.

PICTALKY is the first symbol-based AAC system with neural GEC to provide more accurate and sophisticated language education and communication. PICTALKY automatically outputs the sequence of the pictograms according to the spoken sentences. It can be used for communication between people with disabilities as well as between people with disabilities and non-disabled people. Moreover, it offers the potential to be extended to multilingual versions by using neural machine translation.

2.3 Symbolic AAC

AAC enables nonverbal communication instead of a language, and it can provide practical help for people with cognitive and linguistic disorders.

In the majority of studies on AAC, researchers have employed graphic symbols (i.e., pictograms and picture communication symbols) (Kang et al., 2019) as alternative means of language items to improve the communication skills of children with language developmental disabilities. In this manner, children can be taught how to express their needs and interact with others using symbols (Huang and Lin, 2019).

Most authors have claimed that graphic symbols can enhance the literacy skills and communication of children or support children with disabilities in functional competence (e.g., writing, improving their communication partner knowledge, and learning) (Karal et al., 2016; Nam et al., 2018; Light et al., 2019). Finally, AAC software is a form of symbolic knowledge representation (Beukelman and Mirenda, 2013). That is, symbols are verbal or visual representations of ideas and concepts. Therefore, we adopt both text and image processing mechanisms (i.e., TP) to consider symbolic knowledge with NLP in AAC. Furthermore, we use a deep learning architecture approach for our GEC module. To the best of our knowledge, no such method for a neural and symbol mechanism in AAC has yet been presented.

3 PICTALKY

3.1 Communication Module

Our proposed service uses deep learning-based STT, which takes the voice of the user as input and converts it into text. We adopt Naver CLOVA Speech (Chung, 2019) for the STT system. Furthermore, the text input can be entered with the keyboard as well as in the form of voice. Users

[6]http://www.grembe.com/

and caregivers can enter the text input easily with the keyboards of their personal computer, tablet, or mobile phone.

3.2 Neural GEC Module

People with language disabilities tend to make grammar and pronunciation errors when speaking. The GEC system revises various linguistic errors of users, so it is useful for children to practice correct sentences.

PicTalky is equipped with a neural GEC module that accurately corrects the STT outputs. We denote the sequence-to-sequence model that is applied to the GEC task as neural GEC. From the perspective of machine translation, the neural GEC task is a system whereby a sentence with noise and a correct sentence are entered as the source and target sentences, respectively. Subsequently, translation from the input to the output is trained with the sequence-to-sequence model. In this method, training is conducted without specifying a particular error type; thus, various errors can be detected and processed simultaneously.

PICTALKY enhances the software quality with the latest GEC technique. As a result, the speech errors of people with developmental disorders can be corrected on the text level.

3.3 TP Module

Pictograms are complementary and alternative means of communication that can help people with language difficulties. Unlike languages, which require an understanding of rules and symbolic systems, pictograms deliver the meaning more intuitively and rapidly. Thus, pictograms are utilized in the language rehabilitation field. For example, by using pictograms on communication boards, children can learn how to communicate with others (Calculator and Luchko, 1983). Pictograms provide children who have not learned the language system with practical help in language comprehension and speaking.

This study presents a system that causes the output of the pictogram images to correspond to the input text by using text and image processing. The TP module is an N-gram base mapping system, and it returns the output images that are morphologically similar to the input text in the pictogram dataset. The pictogram dataset includes texts such as words, phrases or sentences that explain the corresponding images. For more accurate mapping, our TP module makes use of a method that scans the entire sentence by N-gram to 1-gram and provides the most similar image.

3.4 NLU module

The output of the TP module is processed by the NLU module to handle the out-of-vocabulary(OOV) text that is not in the pictogram dataset. For this reason, we propose a method that causes the input vocabulary to correspond to a semantically similar image.

In the NLU module, unknown words are replaced with substitute words by measuring the semantic similarities, and a co-reference resolution system is applied to the substitute words. The semantic similarities are measured by Word2Vec (Mikolov et al., 2013) and WordNet (Miller, 1995). Within the input text, substitute words can be resolved through the co-reference resolution function of the spaCy[7] library. The remaining grammatical elements, such as unknown vocabularies, conjunctions, and articles that are not processed by measuring the semantic similarity and replacing unknown words with substitute words are designed not to be printed in the output image.

3.5 Overall Architecture of PICTALKY

When voice input is entered, it is converted into text by the communication module. Subsequently, the text is corrected by the neural GEC system and the corrected texts are changed into pictograms using the TP module. If OOV text exists in the input, the NLU module addresses this problem. Finally, a corresponding pictogram sequence is output.

The overall structure of our proposed service is depicted in Figure 1. If an error sentence `"I lovedd BTS"` is entered as input, the neural GEC corrects the input to `"I love BTS."` Eventually, the text from the pictogram is generated and this module is provided to a form of Web service or robotics.

PICTALKY is aimed at helping children with developmental disabilities to communicate and improve their language understanding. The simultaneous encoding and transmission of speech text, both audibly and visually, allows users to understand the speaker intentions intuitively, even if they have difficulties in using language. Furthermore, as the text and images are delivered together, implicit learning is possible for learning a language by reasoning, without directly teaching each element of the lan-

[7] https://spacy.io/

guage. Thus, the proposed service is intended for children with developmental disabilities, but it can also be applied to rehabilitation for educationally disadvantaged groups.

3.6 PICTALKY with Web Application

We released the PICTALKY as the form of a web application as shown in Figure 2. Thus, any devices enable our system by responsive user interface. The neural GEC module was connected by Rest API and distributed as both CPU and GPU services. Our system is operated by Flask under a cloud server.

Also, we provide the user input into two modes of both speech and text considering the environment of unable to speak. For reviewing other situations on the language disabled, these settings are available to people who have deaf-mutism, aphasia, which is related to being unable to speak. The system built on the compact user interface and freely available to advance accessibility.

Overall procedures of the system are illustrated in Section 3.5. The voice recording starts when the user clicks the record button, and our system begins to print out the result.

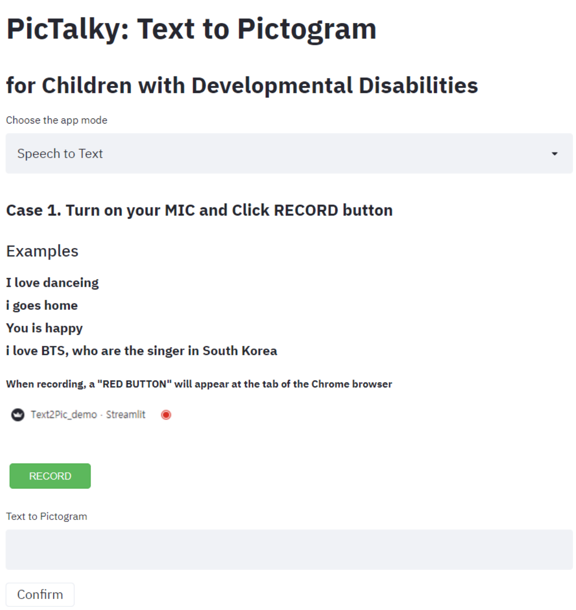


Figure 2: PICTALKY Web application.

4 Experiment and Results

4.1 Datasets

To validate the performance of PICTALKY qualitatively, we adopted a test set that was provided by a GEC service company[8]. The test set was constructed while performing the actual GEC service, inspired by cases in which people with language developmental disabilities utter grammatically incorrect sentences. Thus, it can be stated that it provides high objectivity and reliability. We refer to this test set as the in-house test set. The test set consisted of 100 sentences.

We used parallel corpora as the training data for training our neural GEC model, which were provided by Lang8 (Cho, 2013). We utilized an open-source pictogram dataset that was released by the Aragonese Centre for Augmentative & Alternative Communication[9].

4.2 Verification of Neural GEC Module

Model Although the majority of recent NLP studies have been conducted based on the pretrain-finetuning approach (PFA), it is difficult to service a PFA-based NLP application owing to its slow speed and high computational cost, among other factors (Park et al., 2021). Although state-of-the-art neural models such as mBART (Liu et al., 2020) have been developed, the parameters and model sizes are too large to service in the industry. To overcome this problem, we produced a model based on the vanilla transformer, which is easy to service. The hyperparameters were set to the same values as the settings in Vaswani et al. (2017). The vocabulary size was 32,000 and sentencepiece (Kudo and Richardson, 2018) was adopted for the subword tokenization.

Performance of Neural GEC We used GLEU (Napoles et al., 2015) and BiLingual Evaluation Understudy (BLEU) (Papineni et al., 2002) as evaluation metrics to verify the performance of the neural GEC module. GLEU is similar to BLEU, but it is a more specialized metric for the error correction system, as it considers the source sentences. The overall comparison results are presented in Table 1.

Test set	BLEU	GLEU
In-house	63.77	53.99

Table 1: Performance of neural GEC module.

The experimental results demonstrated that BLEU and GLEU scored 63.77 and 53.99, respectively. These results are sufficiently competitive with the results of other neural GEC studies (Im

[8] https://www.llsollu.com/

[9] http://arasaac.org

	Case	Deletion setting		Score
		POS	Stopwords	
TPA	(1)	✓	✓	**94.16**
	(2)	-	✓	63.96
	(3)	✓	-	52.77
	(4)	-	-	43.59
TPA WITH PENALTY	(1)	✓	✓	**91.62**
	(2)	-	✓	62.24
	(3)	✓	-	51.35
	(4)	-	-	42.42

Table 2: Experimental results of PICTALKY. POS represents the removal of determiners, prepositions, and conjunctions using POS tagging information.

Algorithm 1 TPA

Initialize S_{pos} = {determiner, preposition, conjunction}
`/* The set of exceptional POS tags */`
Initialize S_{stop} as stopwords predefined by NLTK
procedure TPA($sentence$)
  Initialize $score$ and N as zeros
  $W \leftarrow PosTagger(sentence)$
`/* Split words with POS tags */`
  for each word $w \in W$ **do**
    if $w.pos \notin S_{pos}$ and $w \notin S_{stop}$ **then**
      $score \leftarrow score + \delta_{\hat{y},y}$ where $\hat{y} = M_\theta(w)$
`/* Kronecker delta of TP prediction */`
      $score \leftarrow score - (1 - \delta_{\hat{z},z})$ where $\hat{z} = N_\phi(w)$
`/* Penalty for a misclassified named entity */`
      $N \leftarrow N + 1$
  return $score/(N + \epsilon)$

et al., 2017; Choe et al., 2019; Park et al., 2020a,b). This means that our neural GEC module can correct the errors from the STT module, as well as the speech errors of users.

4.3 Verification of TP Module

The results of the performance evaluation of the TP module, which is a core function of PICTALKY, are presented in this section.

TPA We propose TPA, which is a novel metric for measuring the performance of the TP module.

TPA is an objective indicator of how effectively the text in PICTALKY input is converted into pictograms. The measurements are performed as follows. First, the input sentences are separated into words and POS tagged. Thereafter, the words that are POS tagged as determiners, prepositions, conjunctions (POS), and stopwords (Stopwords) are removed, as we believe that these words are meaningless to be converted into pictograms. Thus, the words that do not contain important contents are removed during this process. The remaining words are used for the measurements and the ratio of the words that are effectively converted into pictograms is used as the TPA value. A named entity recognition (NER) penalty is also implemented when calculating the TPA value. The NER penalty is assigned when the named entities are misclassified by the NER process for the input sentences. As the named entities are important information that should be converted without errors, the NER penalty is assigned in those cases. The pseudo-code for the TPA is presented in Algorithm 1.

Case Study According to the deletion setting, we conducted comparative experiments on the TPA with various cases, as indicated in Table 2. There were four cases in total for the deletion setting: (1) both POS (words tagged as determiners, prepositions and conjunctions) and Stopwords are deleted, (2) only Stopwords are deleted, (3) only POS are deleted, and (4) neither POS nor Stopwords are deleted. We also measured how the penalty affected the overall performance. We used NLTK (Bird, 2006) to remove the determiners, prepositions, conjunctions, and stopwords and used the BERT-based (Devlin et al., 2018) NER model provided by Huggingface (Wolf et al., 2019) for the penalty.

The experimental results demonstrated that case (1) of the TPA, which was our proposed method, achieved the highest score of 94.16. In case (2) of the TPA, the score decreased by 30.20 points. When words that were POS tagged as determiners, prepositions, and conjunctions were deleted in case (3), lower performance was exhibited than in case (2). Finally, case (4) achieved the lowest performance. These results demonstrate that excluding both the POS and Stopwords from the subjects of the measurements is the most reasonable evaluation for TP conversion. Moreover, when the NER penalty was applied, the performances decreased in all cases, which means that the NER penalty contributes to more valid measurement.

4.4 Qualitative Analysis

We also conducted a qualitative analysis on the results of PICTALKY based on the developmental stages of the First Language Acquisition (FLA) (Lightbown and Spada, 2021).

In Table 3, the input sentences contain grammatical problems, including fronting, infinitive, article, spelling, plural -s, and irregular past form errors. The PICTALKY Web application shows users the most appropriate pictograms and the output sentences with the errors corrected.

First, `"I love play the baseball"`,

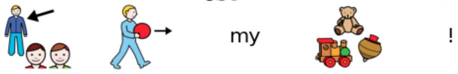
Input sentence	Output sentence	Pictogram
* Is the dog is tired?	Is the dog tired?	
* Do I can eat a pizza?	Can I eat a pizza?	
* I love play the baseball	I love to play baseball	
* I love danceing with a friends	I love dancing with friends	
* He taked my toy!	He took my toy!	

Table 3: Example sentences and pictograms for qualitative analysis created by PICTALKY Web demo.

"I love danceing with a friends", and "He taked my toy!" occur in telegraphic speech (Chomsky, 1964) in the immature language development stage between the ages of two and three. Grammatical errors for morphemes are not merely an imperfect imitation of adults' speech, and consistency of correction with frequent interactions is required to expand cognitive development.

Second, "Is the dog is tired?" and "Do I can eat a pizza?" are one of the errors encountered in acquiring basic structures of the first language between the age of 4 and the school years, and this stage requires correction of low frequency and complex systems. Therefore, we reproduce humans' universal language acquisition process, including frequent errors in the early and later development stages.

Note that if the Neural GEC module cannot correct grammatical errors, the NLU module can compensate for it. However, these aspects need to be supplemented through future research.

5 PICTALKY Satisfaction Survey

We conducted a satisfaction survey to investigate the user satisfaction. It is difficult to employ the nonverbal child. That is why we identified the extreme difficulty in performing a large-scale survey. Therefore, we conducted a system satisfaction survey to 53 people with 43 experts in language disabilities, including ten nonverbal children. The experts consist of thirty teachers of nonverbal children and the thirteen professionals who majored in language disabilities from Korea University Anam Hospital. PICTALKY Satisfaction Questionnaires are shown in Table 4.

We established a total of five questions and specified the answers using a Likert scale (Likert, 1932) of "Satisfied," "Neither agree nor disagree," and "Dissatisfied.". The survey results are depicted in Figure 3.

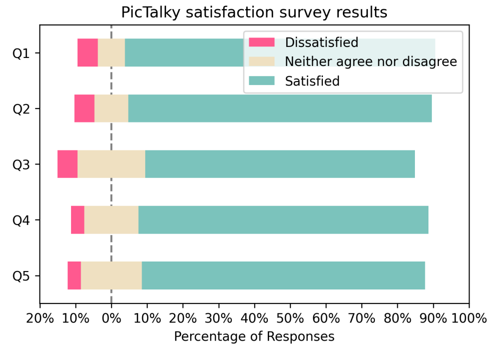


Figure 3: Response results of satisfaction survey regarding PICTALKY.

The survey results revealed that most people were satisfied with the performance of PICTALKY. For each question, 80% to 90% of the responses were satisfied and approximately 90% of the responses stated that it will be helpful to people with developmental disabilities. However, the UI of PICTALKY still requires improvement and the performance of the GEC system should be enhanced. In particular, according to the results of the Spearman correlation (de Winter et al., 2016) of the sentences, as illustrated in Figure 4, the correlation between Q1 and Q2 was high, which indicates that the purpose of this study was well reflected. Although the correlation between Q1 and Q5, and that between Q2 and Q5 were lower than the others, their p-values were lower than 0.05; thus, the results were statistically significant.

Question
Q1. Are you satisfied with the overall performance of PICTALKY?
Q2. Do you think this system will be helpful to people with language developmental disabilities?
Q3. Are you satisfied with the usability and UI of PICTALKY?
Q4. Are you satisfied with the performance of the grammar error correction system?
Q5. Are you satisfied with the results of the text-to-pictogram function?

Table 4: Questions of PICTALKY satisfaction survey.

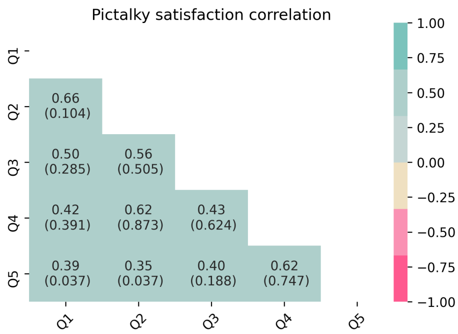


Figure 4: Results of statistical significance test using Spearman correlation between questionnaires. The weight indicates the correlation value and the value in parentheses is the p-value (p-value<0.05 indicates statistical significance).

6 PICTALKY with Robotics

We have distributed PICTALKY as a Web application. However, in the case of the web application, there is a possibility that it is difficult or boring for children to handle.

Therefore, in addition to the Web service, we have applied robotics technology to PICTALKY for arousing interest in children. The NAO robot (Shamsuddin et al., 2011; Jokinen and Wilcock, 2014) is mounted in the communication module of PICTALKY.

NAO is the humanoid robot developed by SoftBank Robotics[10]. Nao has eight full-color RGB LEDs, an inertial sensor, two cameras, and many other sensors. It also has a sonar sensor to check the distance of objects in its vicinity to comprehend its environment with precision and stability. It enables NAO to react its body to move when exposed by interaction. NAO is also available in social robotics (Fong et al., 2003), which focuses on communicating robots capable of interacting and cooperating with humans. All of these characteristics in NAO suit our research pursuit in terms of interacting with a human.

We have created a human–robot interaction system whereby the NAO robot has a conversation with the end users and the pictograms are printed onto the connected screen. As children show substantial interest in robots, this will aid in more familiar education as opposed to Web or other applications (Sennott et al., 2019). The video of our demo is also attached with our paper[11].

To the best of our knowledge, this study is the first to apply PICTALKY to the NAO robot and to develop robotics AAC for the first time.

7 Conclusion and Future Work

We have proposed PICTALKY, which is an AI-based AAC service. The aim of PICTALKY is to provide communication and connection among all people, without anyone being excluded. In the future, we plan to expand the PICTALKY data to multilingual data and to make it fully open. In addition, we will conduct various ai for accessibility studies to improve the quality of life for the disabled.

References

David Beukelman and Pat Mirenda. 2013. Augmentative and alternative communication: Supporting children and adults with complex communication needs. *4th Edition*.

David R Beukelman, Pat Mirenda, et al. 1998. *Augmentative and alternative communication*. Paul H. Brookes Baltimore.

Steven Bird. 2006. Nltk: the natural language toolkit. In *Proceedings of the COLING/ACL 2006 Interactive Presentation Sessions*, pages 69–72.

Stephen Calculator and Christine D'Altilio Luchko. 1983. Evaluating the effectiveness of a communication board training program. *Journal of speech and Hearing Disorders*, 48(2):185–191.

Young Sang Cho. 2013. Lang-8. *Calico Journal*, 30(2):293–299.

Yo Joong Choe, Jiyeon Ham, Kyubyong Park, and Yeoil Yoon. 2019. A neural grammatical error correction system built on better pre-training and sequential transfer learning. *arXiv preprint arXiv:1907.01256*.

[10] https://www.softbankrobotics.com/emea/en/

[11] https://bit.ly/2SunbaW

Noam Chomsky. 1964. [the development of grammar in child language]: Discussion. *Monographs of the Society for Research in Child development*, pages 35–42.

Joon Son Chung. 2019. Naver at activitynet challenge 2019–task b active speaker detection (ava). *arXiv preprint arXiv:1906.10555*.

Joost CF de Winter, Samuel D Gosling, and Jeff Potter. 2016. Comparing the pearson and spearman correlation coefficients across distributions and sample sizes: A tutorial using simulations and empirical data. *Psychological methods*, 21(3):273.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Terrence Fong, Illah Nourbakhsh, and Kerstin Dautenhahn. 2003. A survey of socially interactive robots. *Robotics and autonomous systems*, 42(3-4):143–166.

Chih-Hsiung Huang and Pei-Jung Lin. 2019. Effects of symbol component on the identifying of graphic symbols from eeg for young children with and without developmental delays. *Applied Sciences*, 9(6):1260.

Daniel Im Jiwoong Im, Sungjin Ahn, Roland Memisevic, and Yoshua Bengio. 2017. Denoising criterion for variational auto-encoding framework. In *Thirty-First AAAI Conference on Artificial Intelligence*.

Kristiina Jokinen and Graham Wilcock. 2014. Multimodal open-domain conversations with the nao robot. In *Natural Interaction with Robots, Knowbots and Smartphones*, pages 213–224. Springer.

Rowon Kang, Young Tae Kim, Seok Jeong Yeon, Rowon Kang, Young Tae Kim, and Seok Jeong Yeon. 2019. Cultural differences on the recognition of social word aac graphic symbols between korean and american undergraduate students. *Communication Sciences & Disorders*, 24(1):71–86.

Yasemin Karal, Hasan Karal, Lokman Şılbır, and Taner Altun. 2016. Standardization of a graphic symbol system as an alternative communication tool for turkish. *Journal of Educational Technology & Society*, 19(1):53–66.

Taku Kudo and John Richardson. 2018. Sentencepiece: A simple and language independent subword tokenizer and detokenizer for neural text processing. *arXiv preprint arXiv:1808.06226*.

Eric H Lenneberg. 1967. The biological foundations of language. *Hospital Practice*, 2(12):59–67.

Janice Light, David McNaughton, David Beukelman, Susan Koch Fager, Melanie Fried-Oken, Thomas Jakobs, and Erik Jakobs. 2019. Challenges and opportunities in augmentative and alternative communication: Research and technology development to enhance communication and participation for individuals with complex communication needs. *Augmentative and Alternative Communication*, 35(1):1–12.

Patsy M Lightbown and Nina Spada. 2021. *How Languages Are Learned 5th Edition*. Oxford university press.

Rensis Likert. 1932. A technique for the measurement of attitudes. *Archives of psychology*.

Yinhan Liu, Jiatao Gu, Naman Goyal, Xian Li, Sergey Edunov, Marjan Ghazvininejad, Mike Lewis, and Luke Zettlemoyer. 2020. Multilingual denoising pre-training for neural machine translation. *Transactions of the Association for Computational Linguistics*, 8:726–742.

Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

George A Miller. 1995. Wordnet: a lexical database for english. *Communications of the ACM*, 38(11):39–41.

Sang Nam, Jemma Kim, and Shannon Sparks. 2018. An overview of review studies on effectiveness of major aac systems for individuals with developmental disabilities including autism. *Journal of Special Education Apprenticeship*, 7(2):n2.

Courtney Napoles, Keisuke Sakaguchi, Matt Post, and Joel Tetreault. 2015. Ground truth for grammatical error correction metrics. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 588–593.

Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting on association for computational linguistics*, pages 311–318. Association for Computational Linguistics.

Chanjun Park, Sugyeong Eo, Hyeonseok Moon, and Heui-Seok Lim. 2021. Should we find another model?: Improving neural machine translation performance with one-piece tokenization method without model modification. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Industry Papers*, pages 97–104.

Chanjun Park, Kuekyeng Kim, YeongWook Yang, Minho Kang, and Heuiseok Lim. 2020a. Neural spelling correction: translating incorrect sentences to correct sentences for multimedia. *Multimedia Tools and Applications*, pages 1–18.

Chanjun Park, Yeongwook Yang, Chanhee Lee, and Heuiseok Lim. 2020b. Comparison of the evaluation metrics for neural grammatical error correction with overcorrection. *IEEE Access*, 8:106264–106272.

Steven M Schwarz, Julissa Corredor, Julie Fisher-Medina, Jennifer Cohen, and Simon Rabinowitz. 2001. Diagnosis and treatment of feeding disorders in children with developmental disabilities. *Pediatrics*, 108(3):671–676.

Samuel C Sennott, Linda Akagi, Mary Lee, and Anthony Rhodes. 2019. Aac and artificial intelligence (ai). *Topics in Language Disorders*, 39(4):389–403.

Syamimi Shamsuddin, Luthffi Idzhar Ismail, Hanafiah Yussof, Nur Ismarrubie Zahari, Saiful Bahari, Hafizan Hashim, and Ahmed Jaffar. 2011. Humanoid robot nao: Review of control and motion exploration. In *2011 IEEE international conference on Control System, Computing and Engineering*, pages 511–516. IEEE.

Donghoon Shin, Jaeyoon Song, Seokwoo Song, Jisoo Park, Joonhwan Lee, and Soojin Jun. 2020. Talkingboogie: Collaborative mobile aac system for non-verbal children with developmental disabilities and their caregivers. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–13.

J Bruce Tomblin, Xuyang Zhang, Paula Buckwalter, and Marlea O'Brien. 2003. The stability of primary language disorder. *Journal of Speech, Language, and Hearing Research*.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in neural information processing systems*, pages 5998–6008.

Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, et al. 2019. Huggingface's transformers: State-of-the-art natural language processing. *arXiv preprint arXiv:1910.03771*.

Benjamin Zablotsky, Lindsey I Black, Matthew J Maenner, Laura A Schieve, Melissa L Danielson, Rebecca H Bitsko, Stephen J Blumberg, Michael D Kogan, and Coleen A Boyle. 2019. Prevalence and trends of developmental disabilities among children in the united states: 2009–2017. *Pediatrics*, 144(4).