

CoVoL: A Cooperative Vocabulary Learning Game for Children with Autism

Pawel Chodkiewicz
Department of Computer Science
Reykjavik University
Reykjavik, Iceland
pawel23@ru.is

Pragya Verma
Department of Computer Science
Reykjavik University
Reykjavik, Iceland
pragyav@ru.is

Grischa Liebel
Department of Computer Science
Reykjavik University
Reykjavik, Iceland
grischal@ru.is

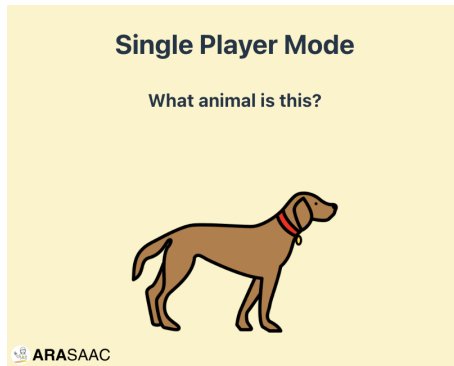


Figure 1: Example prompt for a dog.

Abstract

Children with Autism commonly face difficulties in vocabulary acquisition, which can have an impact on their social communication. Using digital tools for vocabulary learning can prove beneficial for these children, as they can provide a predictable environment and effective individualized feedback. While existing work has explored the use of technology-assisted vocabulary learning for children with Autism, no study has incorporated turn-taking to facilitate learning and use of vocabulary similar to that used in real-world social contexts. To address this gap, we propose the design of a cooperative two-player vocabulary learning game, CoVoL. CoVoL allows children to engage in game-based vocabulary learning useful for real-world social communication scenarios. We discuss our first prototype and its evaluation. Additionally, we present planned features which are based on feedback obtained through ten interviews with researchers and therapists, as well as an evaluation plan for the final release of CoVoL.

CCS Concepts

• **Human-centered computing** → **Interaction design**; • **Applied computing** → **Interactive learning environments**; **Collaborative learning**.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

IDC '25, Reykjavik, Iceland

© 2025 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-1473-3/25/06

<https://doi.org/10.1145/3713043.3731521>

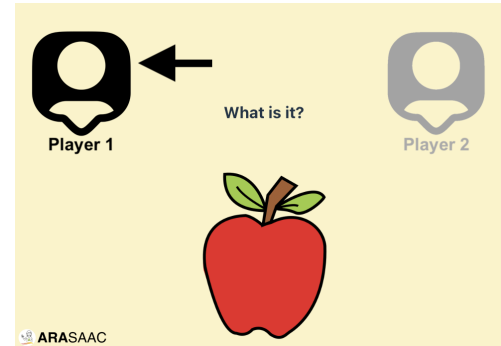


Figure 2: Example prompt for an apple in two-player mode. It is Player 1's turn.

Keywords

Digital Vocabulary Learning, Children with Autism, Turn-taking

ACM Reference Format:

Pawel Chodkiewicz, Pragya Verma, and Grischa Liebel. 2025. CoVoL: A Cooperative Vocabulary Learning Game for Children with Autism. In *Interaction Design and Children (IDC '25)*, June 23–26, 2025, Reykjavik, Iceland. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3713043.3731521>

1 Introduction

Development of language in children takes place in stages and often involves a combination of gestures, vocalizations and use of adequate vocabulary [19]. The acquisition of vocabulary facilitates children's communication in social contexts (known as social communication) and has been shown to be related to academic achievement [16]. Children often are able to comprehend the words spoken to them (receptive vocabulary) before they are able to use them (expressive vocabulary) [13]. In addition, successful social communication requires the ability to link the conversation partner's perspective (Theory of Mind) with the context and the words spoken [5]. Thus, understanding and participating in a social conversation requires one to perceive and comprehend the spoken sounds while considering the conversation partner's perspective (Theory of Mind) and responding using appropriate words.

There is rich literature reporting difficulties in language development of children with Autism [2, 18]. Previous research indicates that children with Autism often face difficulties in vocabulary acquisition that, in turn, can have an impact on their social communication [9]. For teaching vocabulary to children, picture books are often used [10] and have been found to be effective in facilitating vocabulary learning in typically developing children [1].

In comparison to use of traditional picture books, teaching vocabulary using digital tools might have several advantages for children with Autism. For instance, digital tools can offer a more structured and predictable environment, thereby allowing repetition [8]. In addition, it is easier to provide effective individualized feedback [8]. Various researchers have explored the use of technology-assisted vocabulary learning for children with Autism, e.g., [4, 6, 12]. However, none of these past studies have incorporated turn-taking to facilitate learning and use of vocabulary similar to that used in real-world social contexts.

To address this gap, we propose the design of a cooperative (two-player) vocabulary learning game called CoVoL. By combining turn-taking with interactive elements of a digital picture book, children can engage in game-based vocabulary learning useful for real-world social communication scenarios. We believe such an approach will not only facilitate learning and retention of vocabulary, but also improve Theory of Mind in children with Autism by enhancing their understanding of their conversational partner's perspective. This might help children with Autism develop social communication skills, improve their academic performance, and boost their overall confidence and self-esteem.

2 Related Work

In the following, we discuss related work on vocabulary development in children with Autism, as well as work on technological interventions to address existing issues in this group.

2.1 Vocabulary Development in Children with Autism

Up to 50% of children with Autism demonstrate delay in language development [21] that can be attributed to difficulties in various sub-components of language [19]. These difficulties can range from complete absence of verbal ability to difficulties in the use and comprehension of vocabulary [17]. Most vocabulary assessments focus on assessing both expressive and receptive vocabulary [9], both of which children with Autism might experience difficulties with [9]. Some studies report atypical expressive-receptive vocabulary development in children with Autism with expressive vocabulary being larger than receptive vocabulary. Hart et al. [9] find that children with Autism perform poorer than their typically developing counterparts on various aspects of vocabulary starting from 24 months of age. Further, it has been shown that children with Autism are more likely to make semantic errors in addition to being slower in responding to vocabulary tasks [17]. However, despite difficulties in vocabulary acquisition, there seem to be individual trajectories in vocabulary learning for children with Autism [9]. This necessitates the use of individualized vocabulary learning paradigms, as some children may need more intensive vocabulary based programs than others [9]. Such individualized help might be supported by technology.

2.2 Technology-assisted Vocabulary Learning for Children with Autism

Given the benefits associated with technology-based interventions for children with Autism [7] such as increased attention span and

motivation, feasibility of providing automated practice and feedback, possibility of integrating text, sound and images in parallel [22], many researchers in the past have explored the use of technology for vocabulary learning. Specifically, it has been shown that often children with Autism learn from visual media, which can be easily integrated with the help of technology [22]. Previous work has also explored the use of computer-mediated verbal instructions to teach vocabulary to children with Autism [12]. Ganz et al. studied the efficacy of a tablet computer-based augmentative and alternative communication system with voice output on the receptive-language identification of target words [6], showing slight improvements. Chebli et al. used an application to teach one-word receptive identification to children with Autism [4]. In their application, an automated voice was used to name the concept. This was followed by the child choosing the image related to the named concept. Novack et al. used a mobile application that included modified Discrete-Trial Training procedures and other behavioral principals of Applied Behavior Analysis (ABA) to teach receptive vocabulary to children with Autism [14].

However, to the best of our knowledge, no study incorporates peer-mediated learning through turn-taking, which can be effective in enhancing social communication in children with Autism [23]. Additionally, peer-mediated learning might help children in understanding their conversation partner's perspective, i.e., Theory of Mind [5].

3 Method

To implement and evaluate CoVoL, we follow the Design Science Research (DSR) method [3]. In DSR, an artifact is designed and evaluated in an iterative problem-solving process – CoVoL in our case. Each iteration ends with an evaluation step, which leads to novel insights for the next iteration. At the time of submission, we completed the second iteration of the DSR cycle, resulting in a functional prototype of CoVoL for a single player. In the next step, we are extending this to the envisioned two-player game.

3.1 Iteration 1: Designing Activity Flows

In the first iteration, we focused on gaining initial understanding of important design principles prior to implementing a CoVoL prototype. We interviewed three researchers focusing on neurodevelopmental disorders and ABA Verbal Behavior (VB) Therapy based in Poland using a semi-structured interview guide. After providing our initial ideas regarding the features of CoVoL, the experts provided recommendations on activity flows and scenarios suitable for vocabulary learning for children with Autism.

3.2 Iteration 2: Proof of Concept

The second iteration then shifted focus on developing a proof of concept for a single player mode and further investigating core features, so that shifting to the intended two-player mode would be a simple additional implementation step. In parallel, we developed mechanisms to record technical aspects such as word recognition latency and word error rate. To evaluate the second iteration, we tested two English Vosk models with 118 audio samples the first author pre-recorded himself. The 128MB English lgraph model showed an acceptable performance (at over 85% accuracy), but had

a delay of up to 5 seconds on an 2013 MacBook Pro. The 40MB small English model exhibited only around 50% accuracy for the same audio samples. This showed that the larger model is necessary for our use case, and a hosted version might be necessary for slower computers. Additionally, we conducted further interviews with two of the experts from the first iteration and, additionally, with five psychologists or speech-language therapists with clinical experience supporting language development in children with Autism.

3.3 Evaluation Results: Expert Feedback

Based on the input from the expert interviews in Iteration 1 and 2, we elicited various important requirements for CoVoL.

In general, they all stressed the importance of having a free tool for vocabulary learning, ideally with **multi-language support**. In particular, they reported that existing apps are often too expensive for parents to use them at home, and they usually only support English. They also expressed willingness to try out CoVoL in their settings.

The experts supported the implementation of **tacting**, i.e., to tact or label objects [20], as it is the easiest intervention, but crucial in everyday life. They also provided insights into a typical flow, presenting a pictogram every 3 seconds in the beginning, up to 55 pictograms per minute.

In terms of prompts, one expert highlighted that the **same pictogram might be used for different prompts**, not only teaching to label an object, but also attributes such as shape, color or category. This allows children to understand different dimensions.

Rewards were highlighted by several experts. First, they stressed that motivation in children varies substantially, and requires customization options, e.g., to **replace default reward screens**. The timing and frequency of rewards was also stressed in this context. Children get used to overly frequent or fast rewards in electronic applications, thus making it face-to-face human interaction harder. As a consequence, if implemented wrongly, electronic vocabulary learning tools can actually hurt children using them, as they cannot be used together with standard and evidence-based interventions. This highlights the importance of implementing **“human like” and adjustable response times and configuring the frequency of rewards**.

The traditional ABA VB tacting protocol assumes one therapist that prompts a child to label an object., without considering cooperative learning. While the interviewed experts all come from an ABA VB background and were excited about a free game that implements the traditional protocol, they also stressed that the cooperative game idea could be promising. In particular, one expert mentioned that, currently, success in therapy sessions does not always generalize to improved communication outside the therapy. **Cooperation could motivate children to learn social communication** and, thus, allow for a more generalizable learning experience. They highlighted that the two-player game mode models learning in real life, which often takes place through imitation of others' behaviors.

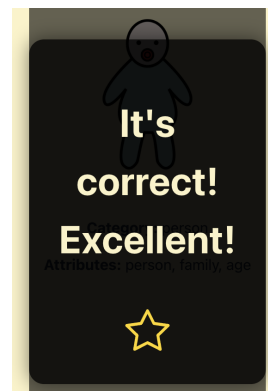


Figure 3: Reward screen for a correctly-labeled object.

4 CoVoL – A Cooperative Vocabulary Learning Game

In the following, we outline the features and technical aspects of CoVoL.

The underlying purpose of CoVoL is to teach children tacting. Improving tacting skills in children with Autism is important, as it can lead to less repetitive and stereotypical language [11]. Following ABA recommendations, in single player mode, we show players a pictogram of an object and prompt them to label this object verbally (see Figure 1). If the player labels the object correctly, a reward screen is shown (see Figure 3). Afterwards, the process is repeated with the next object. This process continues for a defined amount of objects.

In two-player mode, two players take turns labeling objects in the same way as for single player mode. Visually, we display two player avatars on the left-hand side and right-hand side of the screen (see Figure 2). We highlight the avatar who's turn it is, while the other avatar is grayed out. As in single player mode, a reward screen is shown if the object is labeled correctly. Afterwards, it is the other player's turn. In the two player mode, a child learning to speak could be paired with a therapist, a parent/guardian, or another child, potentially depending on their speaking level. In the current prototype of CoVoL, prompts are only displayed textually. Similarly, the players are displayed as “Player 1” and “Player 2” only. In terms of tact recognition, we recognize several synonyms for each object. For instance, the apple prompt in Figure 2 would recognize if the user says “apple”, “apples” or “fruit”.

CoVoL uses a simple three-tier architecture. The frontend is implemented in TypeScript using the Vite framework¹. The backend is implemented in Go using the Fiber framework² and Vosk³ as a voice recognition library. Communication between the frontend and backend takes place through WebSockets. Prompts are stored in JSON files so that non-technical users could easily extend or modify the objects used for tacting. Similarly, object images are stored as

¹<https://vite.dev/>

²<https://gofiber.io/>

³<https://alphacephei.com/vosk/>

regular image files. We use the Augmentative and alternative communication (AAC) image set provided⁴ by the Aragonese Center of Augmentative and Alternative Communication (ARASAAC).

4.1 Planned Features

Based on the obtained expert feedback, we envision various features for CoVoL in addition to the currently implemented aspects. First, we will enable the use of audio prompts and feedback instead (or in addition to) the textual prompts. Additionally, we will allow players to define their own names instead of “Player 1” and “Player 2”, as well as choose an avatar among several options instead of the icons currently displayed (see Figure 2).

Supporting therapists and parents, we will expose various configuration options to the end users. This includes setting a **minimum delay** for the reward, so that children get used to waiting a short while until they receive feedback. Furthermore, we will allow **changing the reward screen**, to tailor it to children’s preferences. For instance, the star could be replaced with another icon. We will implement **multi-language support** in CoVoL. Since Vosk already supports a multitude of languages, this change requires only added support to translate textual prompts and a configuration option to change the language. However, we are also designing the underlying architecture so that Vosk could be exchanged for other language models. Specifically, we are investigating to use an Icelandic language model, tailored to the local context.

Going beyond the basic tacting intervention, we also plan to support different labeling flows. That is, prompts could be changed to ask for attributes of the displayed objects. For example, a player could identify the apple in Figure 2 as “red” or “round”. The data structure used to define the various objects already supports adding attributes, as well as alternative object descriptions (such as “fruit” for the apple). We also envision to support flows that cannot be emulated by a human therapist, such as playing sounds and prompting children to label these.

4.2 Planned Evaluation

In this prototype stage, we only evaluated CoVoL through ten expert interviews. While the author team has experience conducting research in neurodiversity, including children with Autism, and have family members with Autism, any form of evaluation or testing with the intended target group is currently missing. To address this, we are planning to evaluate CoVoL in three scenarios, subject to ethics approval which we are preparing at the moment. First, we will pair one child each with a therapist and evaluate the two-player mode. That is, the therapist and child will take turns labeling objects. Second, we will perform the same kind of evaluation, but pairing a child with one of their parents/guardians instead. In particular, we are interested in how CoVoL can help parents in supporting their children, thus empowering them and reducing the need for therapy sessions. Third, we aim to conduct an exploratory evaluation in which we will pair two children. We will explore both pairs where both children are diagnosed with ASD and pairs where only one child is diagnosed with ASD. Therapists or parents/guardians

will be included as observers in this scenario. We will observe and qualitatively analyze all interactions, as well as following up with interviews with therapists/guardians. These three evaluation scenarios also mirror the intended usage of CoVoL, depending on the child’s speaking aptitude.

4.3 Limitations

We are currently aware of three limitations in CoVoL. First, stuttering can lead to difficulties in detecting words. As stuttering is more common in people with Autism than without [15], this could pose an important barrier to using CoVoL in the intended target group. We will evaluate the extent of this limitation in future work. Second, we currently detect labels on a word level, i.e., whether the correct word has been uttered by the player. This means that we cannot detect negations, as a sentence like “This is not an apple.” would lead Vosk to detect that “apple” has been uttered, thus determining that the object has been correctly labeled. We believe this is not a strong limitation, as we focus on fundamental vocabulary learning only. Finally, we evaluated the quality of Vosk’s word detection with a single person. Further evaluations are necessary to better understand how various dialects or other differences in speaking affect detection quality.

5 Conclusion

We present ongoing work on CoVoL, a two-player game for collaborative vocabulary learning in children with Autism. We designed a prototype in close interaction with researchers in neurodevelopmental disorders and therapists with clinical experience working with children with Autism. The current prototype supports single player mode mirroring the traditional interaction with a therapist. In the next step, we will extend this to a two-player setup with turn taking and evaluate this setup with children. The expert interviews provide rich insights into critical requirements for the design of CoVoL, such as configurable delays to mimic real-life delays in social interaction.

References

- [1] Melissa L Allen, Calum Hartley, and Kate Cain. 2015. Do iPads promote symbolic understanding and word learning in children with autism? *Frontiers in psychology* 6 (2015), 138.
- [2] Vanessa H Bal, Megan Fok, Catherine Lord, Isabel M Smith, Pat Mirenda, Peter Szatmari, Tracy Vaillancourt, Joanne Volden, Charlotte Waddell, Lonnie Zwaigenbaum, et al. 2020. Predictors of longer-term development of expressive language in two independent longitudinal cohorts of language-delayed preschoolers with autism spectrum disorder. *Journal of Child Psychology and Psychiatry* 61, 7 (2020), 826–835.
- [3] Jan vom Brocke, Alan Hevner, and Alexander Maedche. 2020. *Introduction to Design Science Research*. Springer, 1–13. doi:10.1007/978-3-030-46781-4_1
- [4] Sabine Saade Chebli, Marc J Lanovaz, and Marie-Michèle Dufour. 2017. Generalization following tablet-based instruction in children with autism spectrum disorders. *Journal of special education technology* 32, 2 (2017), 70–79.
- [5] Hsiao Yun Chin and Vera Bernard-Opitz. 2000. Teaching conversational skills to children with autism: Effect on the development of a theory of mind. *Journal of autism and developmental disorders* 30 (2000), 569–583.
- [6] Jennifer B Ganz, Ee Rea Hong, Fara Goodwyn, Elizabeth Kite, and Whitney Gilliland. 2015. Impact of PECS tablet computer app on receptive identification of pictures given a verbal stimulus. *Developmental neurorehabilitation* 18, 2 (2015), 82–87.
- [7] Uriel Grynspan, Patrice L Weiss, Fernando Perez-Diaz, and Eynat Gal. 2014. Innovative technology-based interventions for autism spectrum disorders: a meta-analysis. *Autism* 18, 4 (2014), 346–361.
- [8] Arzu Guneyssu, Helena Isabel Silva Reis, Sanna Kuoppamäki, and Cristina Maria Sylla. 2024. Enhancing Autism Therapy through Smart Tangible-Based Digital

⁴The pictographic symbols used are the property of the Government of Aragón and have been created by Sergio Palao for ARASAAC (<http://www.arasaac.org>), that distributes them under Creative Commons License BY-NC-SA.

- Storytelling: Co-Design of Activities and Feasibility Study. In *Proceedings of the 23rd Annual ACM Interaction Design and Children Conference*. 665–669.
- [9] Chelsie M Hart and Suzanne Curtin. 2023. Trajectories of vocabulary development in children with autism spectrum disorder across multiple measures. *Journal of Autism and Developmental Disorders* 53, 4 (2023), 1347–1362.
- [10] Arash Hashemifardnia, Ehsan Namaziandost, and F Rahimi Esfahani. 2018. The effect of teaching picture-books on elementary EFL learners' vocabulary learning. *JELTL (Journal of English Language Teaching and Linguistics)* 3, 3 (2018), 247–258.
- [11] Irfa Karmali, R Douglas Greer, Robin Nuzzolo-Gomez, Denise E Ross, and Celestina Rivera-Valdes. 2005. Reducing palilalia by presenting tact corrections to young children with autism. *The Analysis of Verbal Behavior* 21 (2005), 145–153.
- [12] Kamran Khowaja and Siti Salwah Salim. 2019. Serious game for children with autism to learn vocabulary: an experimental evaluation. *International journal of human-computer interaction* 35, 1 (2019), 1–26.
- [13] Jena McDaniel, Paul Yoder, Tiffany Woynaroski, and Linda R Watson. 2018. Predicting receptive-expressive vocabulary discrepancies in preschool children with autism spectrum disorder. *Journal of Speech, Language, and Hearing Research* 61, 6 (2018), 1426–1439.
- [14] Marlena N Novack, Esther Hong, Dennis R Dixon, and Doreen Granpeesheh. 2019. An evaluation of a mobile application designed to teach receptive language skills to children with autism spectrum disorder. *Behavior analysis in practice* 12 (2019), 66–77.
- [15] Veera Pirinen, Soile Loukusa, Katja Dindar, Leena Mäkinen, Tuula Hurtig, Katja Jussila, Marja-Leena Mattila, and Kurt Eggers. 2023. A Comprehensive Analysis of Speech Disfluencies in Autistic Young Adults and Control Young Adults: Group Differences in Typical, Stuttering-Like, and Atypical Disfluencies. *Journal of Speech, Language, and Hearing Research* 66, 3 (2023), 832–848. doi:10.1044/2022_JSLHR-22-00265
- [16] K Ashana Ramsook, Janet A Welsh, and Karen L Bierman. 2020. What you say, and how you say it: Preschoolers' growth in vocabulary and communication skills differentially predict kindergarten academic achievement and self-regulation. *Social Development* 29, 3 (2020), 783–800.
- [17] David Saldaña. 2023. Atypical vocabulary acquisition in autism: where is it coming from? *Journal of Cultural Cognitive Science* 7, 1 (2023), 1–7.
- [18] Joanne Elizabeth Saul. 2020. *Expressive language development in minimally verbal autistic children: exploring the role of speech production*. Ph. D. Dissertation. UCL (University College London).
- [19] H Shilpashri and K Shyamala. 2020. Pragmatic Skills during Mother-Child Interaction in Children with Autism. *Int. J. Health Sci. Res* 10 (2020), 244–250.
- [20] Burrhus Frederic Skinner. 1957. *Verbal behavior*. Appleton-century-crofts.
- [21] Veronica Smith, Pat Mirenda, and Anat Zaidman-Zait. 2007. Predictors of expressive vocabulary growth in children with autism. *Journal of Speech, Language, and Hearing Research* 50, 1 (2007), 149–160.
- [22] Ana Lucia Urrea, Vanessa Fernández-Torres, Isabel R Rodriguez-Ortiz, and David Saldaña. 2024. The use of technology-assisted intervention in vocabulary learning for children with autism spectrum disorder: a systematic review. *Frontiers in Psychology* 15 (2024), 1370965.
- [23] Beihua Zhang, Shan Liang, Jingze Chen, Lin Chen, Weimin Chen, Shunshun Tu, Linyan Hu, Huimin Jin, and Lixi Chu. 2022. Effectiveness of peer-mediated intervention on social skills for children with autism spectrum disorder: a randomized controlled trial. *Translational Pediatrics* 11, 5 (2022), 663.