Journal ofIntelligent Communication (2023) Volume 2 Issue 1 , 9 *pages*
.doi: 10.54963/jic.v2i2.127

Original Research Article

Design of an Intelligent Agent to Measure Collaboration and Verbal-Communication Skills of Children with Autism Spectrum Disorder in Collaborative Puzzle Games

Miao Zhang[1*], Chengwei Hua[2]

*[*1] Vanderbilt University, Nashville, TN, USA. Email: MZhang@nyu.edu*

[2] China Transport Telecommunications & Information Center, Beijing 100011, China

ABSTRACT

Autism Spectrum Disorder (ASD) is aneurodevelopmental disorder characterized by core deficits in social interaction and communication. Collaborative puzzle games are interactive activities that can be played to foster the collaboration and verbal-communication skills of children with ASD. In this paper, we have designed an intelligent agent that can play collaborative puzzle games with children and verbally communicate with them as if it is another human player. Furthermore, this intelligent agent is also able to automatically measure children's task-performance and verbal-communication behaviors throughout game play. Two preliminary studies were conducted with children with ASD to evaluate the feasibility and performance of the intelligent agent. Results of Study I demonstrated the intelligent agent's ability to play games and communicate with children within the game-playing domain. Results of Study II indicated its potential to measure the communication and collaboration skills of human users.

Keywords: Traffic accidents; Communication tools; Power-minimization

1. Introduction

Autism Spectrum Disorder (ASD) is aneurodevelopmental disorder characterized by core deficits in social interaction and communication[1]. The estimated prevalence of ASD in the United States is 1 in 59, as reported by the Centers for Disease Control and Prevention[2]. The individual incremental lifetime cost associated with ASD is over $3.2 million[3]. With its high prevalence rate and associated costs, a wide range of studies have explored mechanisms to positively impact the social communications of children with ASD as well as the improvement of their long-term developmental outcomes[4,5]. Although a cumulative literature review suggests that some

ARTICLE INFO

Received: August 5, 2022 | Accepted: September 29, 2022 | Available online: October 12, 2022

CITATION

Miao Zhang, Chengwei Hua. Design of an Intelligent Agent to Measure Collaboration and Ver-bal-Communication Skills of Children with Autism Spectrum Dis-order in Collaborative Puzzle Games, Journal of Intelligent Communication, 2022; 2(2): 12–20 pages.

interventions can have positive impacts on the lives of children with ASD and their families, many families struggle to access evidence-based care due to its high cost (often over $100/hour, with recommended intensity of at least 15 hours per week) and a shortage of trained clinicians [6–15]. Therefore, an urgent need exists for inexpensive, accessible, and effective assistive therapeutic modalities for ASD intervention.

Computer-assisted interventions may offer an alternative intervention and assessment modality with reduced costs of care. Many children with ASD have a natural affinity for computer-controlled environments and exhibit a high level of engagement within these systems. In addition, computer systems can provide controllable, replicable, and safe environments for children with ASD to practice social communication skills. As such, various kinds of computer-mediated intervention systems have been developed in order to understand and enhance the social communication skills of children with ASD. Among these are collaborative game-based interventions, which usually target two users' ability to convey information to one another (communicate) and to work together to achieve a common goal (collaborate).

Collaborative games poses several advantages relative to traditional intervention and assessment modalities. First, many children with ASD show a high level of engagement in computer-based collaborative games. Hourcade and colleagues designed four computer games that required two users to work together. They analyzed users' collaborative interactions by manually coding the users' conversations within the system, and found that children with ASD spoke more sentences when they played these computer games as compared to non-computer games. Another advantage of collaborative computer games is that they can be designed to include strategies to elicit collaborative skills. Battocchi and colleagues designed collaborative puzzle games with an enforced collaboration rule, which required two users to take actions simultaneously, in order to encourage collaborations between the users. They evaluated the effect of these games on users' collaborations by measuring users' task performance, such as task completion time and number of moved puzzle pieces. They found that these collaborative games, equipped with the enforced collaboration rule, have more positive effects on children with ASD, compared to games without such rules. Piper and colleagues designed and implemented a cooperative tabletop computer games for adolescents with Asperger's Syndrome. They found that the cooperative computer games improved engagement, group work skills and confidence to interact in social activities within the population. Other previous literature on collaborative games with intervention strategies have also successfully investigated other collaborative behaviors of children with ASD, such as sharing[16], turn taking[17], and collaborative play[18].

In this work, we present the design of an intelligent agent able to play collaborative games with children with ASD, and simultaneously communicate with them during the games. In addition, the intelligent agent was designed to automatically measure collaboration and verbal-communication skills of children with ASD when they played these games. Such an intelligent agent may have the ability to 1) encourage collaborative interaction and communication in children with ASD; and 2) automatically evaluate the impacts of these collaborative games on these children. We also conducted two preliminary studies to evaluate the feasibility of the intelligent agent to interact with the target population, as well as its potentialto measure their collaboration and verbal-communication skills.

The main challenge of designing such an intelligent agent is to understand the unrestricted human language using a computer program. Note that designing a computer program that can understand human language and conduct conversations as a human (i.e., Turing test) is yet to be solved from a technical point of view. Existing intelligent agents with conversation capabilities can only work in narrowly defined domains [19–22]. In our implementation, the intelligent agent was also designed with narrowly defined domains when communicating and playing games with children with ASD.

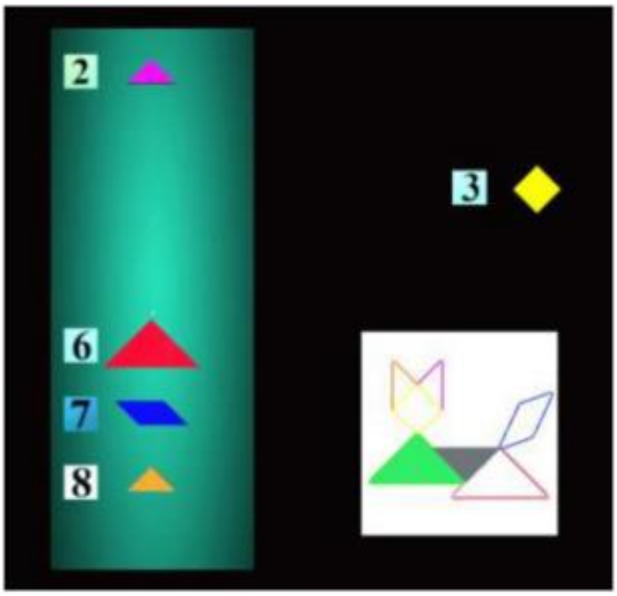

Figure 1: An example of the collaborative puzzle games.

2. Intelligent Agent

A. Overall Description and Architecture

We designed an intelligent agent with the ability to elicit and assess COllaboratioN and COmmunicatioN (ICON2) skills of children with ASD through games and conversation tasks. The overall view of ICON2 is shown in Fig. 2. ICON2 can perceive a human's speech and game-related actions, i.e., what the human-partner says and what he/she does to play collaborative puzzle games. Then, it generates speech and game-related actions based on the perceived information. Finally, it executes these generated speech and game-related actions as responses to the human. ICON2 monitors input in realtime without requiring the presence of a human therapist or coder. As ICON2 plays the games, as described below, it can assess collaborative and communicative aspects of the interaction through machine learning methods using several collaboration and communication features that were defined based on previous work.

A human user interacts with ICON2 when playing collaborative puzzle games. ICON2 acts as a virtual partner that is capable of conversing with the human user and also executes game actions during game play. ICON2 is aware of the game states, the rules of the games, and the layout of the virtual environment. As described in section II, the puzzle games employ two configurations to promote collaboration, "turn-taking" and "move together." When the game is in the "turn-taking" configuration, the human user first moves a puzzle piece to the target image, and ICON2 observes the movement and waits for its own turn. If the human user does not make a move, ICON2 will prompt the user by saying, "This isaturn-taking game. It is your turn to move the puzzle piece." When it is ICON2's turn to move a puzzle piece, it asks the human user which piece it should move. It then "listens" to what the human user says and independently moves the identified piece to the target. When the game is in a "move together" configuration, ICON2 waits for the human user to communicate which piece to move together, verbally confirms the selection, and moves the piece together with the human user. If the human user does not verbally communicate which piece to move (e.g., User silently clicks on a piece to move), ICON2 will attempt to prompt user communication by asking, "Which piece should we move?"

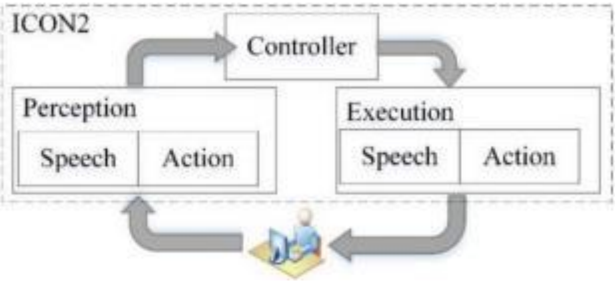

Figure 2: Overall view of ICON2.

A human user interacts with ICON2 when playing collaborative puzzle games. ICON2 acts asavirtual partner that is capable of conversing with the human user and also executes game actions during game play. ICON2 is aware of the game states, the

rules of the games, and the layout of the virtual environment. As described in section II, the puzzle games employ two configurations to promote collaboration, “turn-taking” and “move together.” When the game is in the “turn-taking” configuration, the human user first moves a puzzle piece to the target image, and ICON2 observes the movement and waits for its own turn. If the human user does not make a move, ICON2 will prompt the user by saying, “This is a turn-taking game. It is your turn to move the puzzle piece.” When it is ICON2’s turn to move a puzzle piece, it asks the human user which piece it should move. It then “listens” to what the human user says and independently moves the identified piece to the target. When the game is in a “move together” configuration, ICON2 waits for the human user to communicate which piece to move together, verbally confirms the selection, and moves the piece together with the human user. If the human user does not verbally communicate which piece to move (e.g., User silently clicks on a piece to move), ICON2 will attempt to prompt user communication by asking, “Which piece should we move?”

ICON2 is aware of the human user game actions (puzzle piece locations and mouse clicks locations) and combines this information with verbal input from the human user. For example,a human user can opt to get color information from ICON2 in two ways:

Human user can ask ICON2, “What color is puzzle 5?” or

Human user can use the mouse to click on puzzle 5 and ask ICON2, “What is the color for this?”

ICON2 can correctly respond to both user actions with the color information of puzzle 5, even though in the later way the user did not specify the number of the puzzle piece.

Different game characteristics are employed to encourage communication and collaboration in the human user and described in Table II and Table III. For example, in both configurations, when the color information of the puzzle pieces is only available to one user (human user or ICON2), both partners have to exchange the color information in order to move the pieces to the target image. If prompted by the human user, ICON2 will always respond with the correct color information. When color information is not available for ICON2, it will ask the human user the color information before the puzzle piece can be moved to the target image.

3. Skill Measurements Procedure

We now present the procedure to measure both communication and collaboration skills. The system generated task-performance and verbal-communication features to represent the participants’ behaviors when they interacted with ICON2 in the CVE. Then we applied machine learning methods to measure these skills based on the system-generated features.

A. System-Generated Features

The system automatically generated multiple verbal-communication and task-performance features, which were designed based on previous literature in the field. All the features and their definitions are shown in Table VI. Previous literature demonstrated that dialogue act features, such as requests for information, providing information, and acknowledging other people’s actions, were useful in understanding group discussion behaviors of both children with ASD and TD children. In addition, word frequency and sentence frequency have been found useful to reflect the behaviors of children with ASD during collaborative puzzle games. Bauminger-Zviely and colleagues found that the success frequency and failure frequency reflected important aspects of collaborative behaviors of children with ASD in collaborative puzzle games. White and colleagues reported that the dragging time and collaboration time features could reflect collaborative efficiency of children with ASD when they played collaborative puzzle games with their TD peers. In our system, all the features shown in Table VI were generated by the system in real-time and recorded for offline analysis.

B. Skill Measurements

We built machine learning models to measure participants' communication and collaboration skills using the system-generated features. In particular, we trained machine learning models to classify a data sample, which included all system-generated features of a game, into abinary-class, i.e., a high level of skills or a low level of skills. First, we applied Principal Component Analysis (PCA) to reduce the feature dimension. Then we trained a Support Vector Machine with Radial Basis Function (SVM-RBF) model to measure communication skills using the system-generated features and ratings of communication skills on a binary scale, and trained another SVM-RBF model to measure collaboration skills using the features and rating of the collaboration skills on a binary scale. We selected SVM-RBF kernel as the machine learning method for the classification because this method usually performs well in classifying data with a small sample size. The performance of these models in measuring these skills was evaluated using their classification accuracies, which were computed using a 6-fold cross-valuation method.

4. Conclusions

In this paper, we designed an intelligent agent that could communicate and play games with children with ASD in a CVE as well as generate meaningful features to measure their communication and collaboration skills. Results of the two preliminary studies presented here indicate the potential of ICON2 to 1) communicate and collaborate with children with ASD in the CVE as indicated by the self-report results; and 2) generate meaningful features to measure communication and collaboration skills of the participants as indicated by high accuracies of these measurements.

In particular, we found that ICON2 could appropriately initiate conversations and respond to the participants' conversation in Study I. ICON2 generated 82.93% appropriate initiations and 89.33% appropriate replies when interacting with the children with ASD. These accuracies are comparable to results of other intelligent agents with conversation capabilities designed for TD individuals. Given differences in data sample numbers and task domains, it is hard to directly compare numerical results of different systems in this area. However, we believe that the communication capability of ICON2 are comparable to existing systems by comparing the numerical results available in the literature. For example, Kopp and colleagues designed a conversational agent as a museum guide to communicate with museum visitors. The agent could understand visitors' utterances by mapping keywords with 138 rules. The agent could correctly respond to visitors' 50423 utterances with an accuracy of 63%. Tewari and colleagues designed a question-answer system to help improve reading skills of children in the lowest socio-economic status. The system could correctly answer questions with an accuracy of 86%, which was computed with 346 utterances. However, this system could not initiate conversations and did not support non-speech interactions. Ramin and colleagues designed a spoken system to assist elderly users about their weekly planning. The system could respond to elderly users with 84.8% accuracy, which was computed from only 46 utterances.

ICON2 has the potential to evaluate communication and collaboration skills of the participants as seen in Study II. The system could accurately generate verbal-communication features as indicated by the low error rates of these features. For example, the sentence frequency feature had a low error rate 0.0566. All the features together could measure these skills with high accuracies using machine learning models. The accuracy to measure the communication skills was 89.58%, while the accuracy to measure the collaboration skills was 75.40%. Although these machine learning models were built offline, they could be used for real-time measurements in the future. The results indicate that the system has the potential to automatically measure both communication and collaboration skills in human-agent interactions based on these system-generated features. Automated systems for capturing, labeling, and measuring communicative overtures could, in the future, augment our ability to systematically measure change in important social-communication therapy goals. This has the potential to reduce costs associated with human observation and coding as well as reducing subjective bias in behavioral observation. That being said, in the current work the system required intensive human-coder classification in order to develop and optimize our models. Future, use of such a paradigm will ultimately have to overcome this system development cost to move toward larger scale use.

The errors that occurred when the system generated verbal-communication features were because of errors in speech recognition and errors in dialogue act classification. Errors of the word frequency and the sentence frequency features were due to errors of the speech recognition; while errors of other verbal-communication features, such as the Request Color frequency, Provide frequency, and Direct Movement frequency, were due to both the speech recognition errors and the dialogue act classification errors, as shown in Table X. This might be the reason why the word frequency and sentence frequency features had the lowest error rates. We also found a high error rate for the Request Object frequency feature. This maybe because the participants spoke only a few Request Object sentences, as indicated by the small ratio of the Request Object frequency to the sentence frequency in Table X. As a result, a few incorrectly detected Request Object sentences could lead to a high error rate. We found similar results regarding the Request Color frequency feature.

While we have presented a novel hybrid method to develop this intelligent agent for meaningful measurements within the tangram puzzles domain with varying configurations (colors, no colors, turn taking, move together), ICON2'scommunication behaviors could be extended to other domains by modifying the hybrid method.ICON2's generated speech responses within the game-playing domain based on some rules can be extended to other domains by modifying these rules. After adjusting the variables of the hybrid method, ICON2 will be able to communicate and interact with users in other domains. Also, ICON2 design was not adaptive, where the system performed at the same level for users from different age and developmental group. It would be worth exploring the influence of varying the type of game as well as incorporating different difficulty levels to the communication and collaboration skills of the participants.

Although the present work is promising, readers are advised to exercise caution in interpreting theresults more generally due to several limitations of the current work. First, the sample size was small, and the experimental design consisted of only one session. Please note that the goal of the present study was to design an intelligent agent that could play collaborative games and communicate within the game-playing domain to automatically measure important aspects of interactions in a CVE with preliminary studies. Results of the preliminary studies indicated that this intelligent agent has the potential to interact with children with ASDas well as automatically generate meaningful features to measure both communication and collaboration skills of children with ASD. In the next step, we will utilize this system for real-time measurements with more participants and with a longer study duration.

Second, the use of binary scale (0 = low, 1 = high) to rate the communication and collaboration skills may not be sufficient to provide in depth and continuous measure of these skills. The binary scale was used as an initial step to assess the feasibility of the agent without adding complexity of the analysis. Moving forward, work beyond proof of concept could possibly explore a more refined rating scheme that would be able to provide in depth rating of both skills.

Third, the training data used to build the SVM-RBF model for the dialogue act classification was relatively small. While the accuracy (67.47%) of the classifier in Study I and the accuracy (69.10%) of the classifier in Study II were much higher than the random accuracy (i.e., 20%) of a five-class classifier, more training data may yield a classification model with higher accuracy. In addition, the out-of-domain detection method in this paper was limited. Future studies should aim to develop more efficient methods for out-of-domain detection.

Fourth,the system-generated features were limited as well. We only explored 12 features for the measurements in the current study. Human behaviors, such as eye gaze, body language, and facial expression, could also provide important information in peer-mediated interactions. However, features to represent these behaviors have not been explored in this study. In the future, these features will be captured using eye gaze recognition, gesture recognition, and emotion recognition in order to understand the non-verbal communications. And unlike an actual human partner, ICON2 has the potential to crash or fail. This could cause user frustration. As mentioned earlier in the section, the system did crash and caused data loss for one session, but the system was recovered right away to not cause further disruption to the experiment Despite these limitations, the performance of the games and interactions of the participants with their partners and the system itself were not affected and further contributes to the collaborative learning literature by proposing a novel way to automatically measure communication and collaboration skills of children with ASD within a CVE using an intelligent agent. Results of the two studies indicated that the presented intelligent agent was tolerated and apparently engaging and enjoyable to the participants, as well as demonstrate its potential to automatically measure important aspects of interactions in a CVE. The scope of the current work was to design the intelligent agent and preliminarily assess its capability to capture both communication and collaboration skills of children with ASD when they interacted with the intelligent agent in a CVE. This is a necessary first step before intelligent agents could be strategically deployed to assess these skills during peer-to-peer interactions within similar collaborative environments.

Conflict of interest

The authors declare no conflict of interest.

Funding

This work was supported in part by the National Institute of Health Grants 1R01MH091102-01A1 and 1R21MH111548-01, and National Science Foundation Grant 0967170 and the Hobbs Society Grant from the Vanderbilt Kennedy Center. The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

References

1. Snoek CGM, Worring M, and Smeulders AWM, “Early versus late fusion in semantic video analysis,” in Proc. 13th ACM Int. Conf. Multimedia, MM 2005, 2005, pp. 399–402, doi: 10.1145/1101149.1101236.
2. Baio J et al., “Prevalence of Autism Spectrum Disorder among children aged 8 years-Autism and Developmental Disabilities Monitoring Network,” 2014, MMWR Surveillance Summaries. doi: 10.15585/mmwr.ss6706a1.
3. Peacock G, Amendah D, Ouyang L, and Grosse SD, “Autism Spectrum Disorders and Health Care Expenditures,” J. Dev. Behav. Pediatr, vol. 33, no. 1, pp. 2–8, Jan. 2012, doi: 10.1097/DBP.0b013e31823969de.
4. Sundberg ML and Partington JW, Teaching Language to Children with Autism Or Other Developmental Disabilities, Pleasant Hill, California, USA: Behavior Analysts, Inc., 1998.
5. Bauminger N, “The facilitation of social-emotional understanding and social interaction in high-functioning children with autism: Intervention Outcomes,” J. Autism Dev. Disord, vol. 32, no. 4, pp. 283–298, Aug. 2002, doi: 10.1023/A:1016378718278.
6. Rogers SJ, “Empirically supported comprehensive treatments for young children with autism,” J. Clin. Child Psychol, vol. 27, no. 2, pp. 168–179, 1998, doi: 10.1207/s15374424jccp2702_4.

7. Cohen H, Amerine-Dickens M, and Smith T, "Early intensive behavioral treatment: Replication of the UCLA model in a community setting," J. Dev. Behav. Pediatr, vol. 27, pp. S145–S155, 2006. doi: 10.1097/00004703-200604002-0001.
8. Pennington RC, "Computer-assisted instruction for teaching academic skills to students with Autism Spectrum Disorders: a review of literature," Focus Autism Other Dev. Disabl, vol. 25, no. 4, pp. 239–248, Dec. 2010, doi: 10.1177/1088357610378291.
9. MooreD, Cheng Yufang, Mcgrath P, and Powell NJ, "Collaborative virtual environment technology for people with autism," Focus Autism Other Dev. Disabl, vol. 20, no. 4, pp. 231–243, 2005, doi: 10.1177/10883576050200040501.
10. Lahiri U, Bekele E, Dohrmann E, Warren Z, and Sarkar N, "A physiologically informed virtual reality based social communication system for individuals with autism," J. Autism Dev. Disord, vol. 45, pp. 919–931, 2015, doi: 10.1007/s10803-014-2240-5.
11. Bernard-Opitz V, Sriram N, and Nakhoda-Sapuan S, "Enhancing social problem solving in children with Autism and normal children through computer-assisted instruction," J. Autism Dev. Disord, vol. 31, no. 4, pp. 377–384, Aug. 2001, doi: 10.1023/A:1010660502130.
12. Noor H, Shahbodin F, and Pee NC, "Serious game for autism children: review of literature," World Academy Sci., Eng. and Technol. Int. J. Psychol. and Behav. Sci, vol. 6, pp. 554–559, 2012, doi: 10.5281/zenodo.1333272.
13. Hourcade JP, Williams SR, Miller EA, Huebner KE, and Liang LJ, "Evaluation of tablet apps to encourage social interaction in children with Autism Spectrum Disorders," in Conf. Human Factors Comput. Syst – Proc, 2013, pp. 3197–3206, doi: 10.1145/2470654.2466438.
14. BattocchiA, Pianesi F, Tomasini D, ZancanaroM, Esposito G, Venuti P, Ben Sasson A, Gal E, and WeissPL, "Collaborative puzzle game: Atabletop interactive game for fostering collaboration in children with Autism Spectrum Disorders (ASD)," in Proc. ACM Int. Conf. Interactive Tabletops and Surfaces, 2009, pp. 197–204, doi: 10.1145/1731903.1731940.
15. Piper AM, O'Brien E, Morris MR, and Winograd T, "SIDES: A cooperative tabletop computer game for social skills development," in Proc. ACM Conf.Comput. Supported Cooperative Work, CSCW, 2006, pp. 1–10, doi: 10.1145/1180875.1180877.
16. Curtis DD and Lawson MJ, "Exploring collaborative online learning," J. Asynchronous Learn. Netw, vol. 5, pp. 21–34, 2001, doi: 10.24059/olj.v5i1.1885.
17. ZancanaroM, Pianesi F, Stock O, Venuti P, CappellettiA, Iandolo G, PreteM, and Rossi F, "Children in the museum: an environment for collaborative storytelling," in PEACH-Intell. Interfaces Museum Visits, 2007, pp. 165–184. doi: 10.1007/3-540-68755-6_8.
18. Ben-Sasson A, Lamash L, and Gal E, "To enforce or not to enforce? The use of collaborative interfaces to promote social skills in children with high functioning autism spectrum disorder," Autism, vol. 17, pp. 608–622, 2013, doi: 10.1177/1362361312451526.
19. Kopp S, Gesellensetter L, KrämerNC, and Wachsmuth I, "A conversational agent as museum guide-design and evaluation of a real-world application," in Int. Workshop Intell. Virtual Agents, 2005, pp. 329–343, doi: 10.1007/11550617_28.
20. Cauell J, Bickmore T, Campbell L, and Vilhjálmsson H, "Designing embodied conversational agents," Embodied Conversational Agents,

Cambridge, Massachusetts, USA: MIT Press, 2000, pp. 29–63.

21. Pellom B, Ward W, Hansen J, Cole R, HaciogluK, Zhang J, Yu X, and Pradhan S, "University of Colorado dialog systems for travel and navigation," in Proc.1st Int. Conf. Human Lang. Technol. Res, 2001, pp. 1–6, doi: 10.3115/1072133.1072225.
22. Aust H, Oerder M, Seide F, and Steinbiss V, "The Philips automatic train timetable information system," Speech Commun, vol. 17, pp. 249–262, 1995, doi: 10.1016/0167-6393(95)00028-M.