TRIK: A Talking and Drawing Robot for Children with Communication Disabilities

Peter Ljunglöf
Staffan Larsson
University of Gothenburg
Gothenburg, Sweden

Katarina Mühlenbock
Gunilla Thunberg
DART: Centre for AAC and AT
Gothenburg, Sweden

Abstract

This paper describes an ongoing project where we develop and evaluate setup involving a communication board (for manual sign communication) and a drawing robot, which can communicate with each other via spoken language. The purpose is to help children with severe communication disabilities to learn language, language use and cooperation, in a playful and inspiring way. The communication board speaks and the robot is able to understand and talk back. This encourages the child to use the language and learn to cooperate to reach a common goal, which in this case is to get the robot to draw figures on a paper.

1 Introduction

1.1 Dialogue systems

Most existing dialogue systems are meant to be used by competent language users without physical or cognitive language disabilities – either they are supposed to be spoken to (e.g., phone based systems), or one has to be able to type the utterances (e.g., the interactive agents that can be found on the web). The few dialogue systems which are developed with disabled people in mind are targeted at persons with physical disabilities, who need help in performing common acts.

Dialogue systems have also been used for second language learning; i.e., learning a new language for already language competent people. Two examples are the artificial agent *"Ville – The Virtual Language Tutor"* (Beskow et al., 2004), and *"SCILL – Spoken Conversational Interface for Language Learning"*, a system for practicing Mandarin Chinese (Seneff et al., 2004).

However, we are not aware of any examples where a dialogue system has been used for improving first language learning.

1.2 Target audience

Our intended target group are children with severe communication disabilities, who needs help to learn and practice linguistic communication. One example can be children with autism spectrum disorders, having extensive difficulties with representational thinking and who therefore will have problems in learning linguistic communication. Many children with autism are furthermore hindered in their speech development by the fact that they also have physical disabilities. Our dialogue system will give an opportunity to explore spoken language – content as well as expression.

Another target audience which we believe will benefit from our system are children whose physical disabilities are very extensive, usually as a consequence of Cerebral Palsy (CP). The ablility to control a robot gives a fantastic opportunity to play, draw and express oneself in spoken language, which otherwise would be very difficult or even impossible.

1.3 Language development

To be able to learn a language one must have practice in using it, especially in interplay with other language competent people. For the communication to be as natural as possible, all participants should use the same language. For that reason there is a point in being able to express oneself in spoken language, even if one does not have the physical or cognitive ability. If one usually expresses oneself by pointing at a communication board, it is thus important that the board can express in words what is meant by the pointing act. This is even more important when learning a language, and its expressions and conventions (Sevcik and Romski, 2002; Thunberg, 2007).

When it comes to children with autism, learning

Kristiina Jokinen and Eckhard Bick (Eds.)
NODALIDA 2009 Conference Proceedings, pp. 275–278

appears to be simpler in cooperation with a technical product (e.g., a computer), since the interaction in that case is not as complex as with another human (Heimann and Tjus, 1997). Autistic persons have difficulties in coordinating impressions from several different senses and different focuses of attention. When one is expected to listen to, look at and interpret a number of small signals, all at the same time, such as facial expressions and gazes, human communication can become very difficult.

All children need repetition to learn things. Children with disabilities often need even more repetition in their language learning, because of their lack of communicative functions. Adapted techniques, and in this case the speech-controlled drawing robot, can offer the required repetition as an exciting complement to human communication.

2 Project description

Our basic idea is to use a dialogue system to support language development for children with severe communicative disabilities. There are already communication boards connected to speech synthesis in the form of communication software on computers. The main values that this project add to existing systems are that

1. the child can explore language on her own and in stimulating cooperation with the robot;
2. it can be relieving and stimulating at the same time, with a common focus on the dialogue together with a robot;
3. the child is offered an exciting, creative and fun activity.

By being able to use a picture- or symbol-based communication board the children are given an exciting opportunity to explore language; to play and in the same time learn to use a method for alternative and augmentative communication.

2.1 A talking communication board and a talking robot

In our goal scenario the child has a communication board which can talk; i.e., when the child points at some symbols they are translated to an utterance which the board expresses via speech synthesis, and in grammatically correct Swedish. This is recognized by a robot which can move around on a paper and draw at the same time. The robot executes the commands that was expressed by the communication board; e.g., if the child points at the symbol for *"draw a figure"*, and the symbol with a flower, the utterance might be *"draw a flower, please"*, which the robot then performs.

The dialogue system comes into play when the robot is given too little information. E.g., if the child only points at the symbol for *"draw a figure"*, the robot does not get enough information. This is noticed by the dialogue system and the robot asks a follow-up question, such as *"what figure do you want me to draw?"*.

2.1.1 Functionality of the robot

Our robot is a variant of the LOGO-robot which was developed at Massachusetts Institute of Technology for learning children to use computers and program simple applications (Papert, 1993). The robot can move forward and backward, and turn right and left. It also has a pen which it can lift (for not drawing) or lower (for drawing). The robot can also be programmed to execute command sequences; e.g., it is possible to define that a *"square"* is to first move forward, turn left 90 degrees, and then redo the same thing three more times.

2.2 Pedagogical advantages

By having the communication board and the robot talking to each other there is a possibility for users in an early stage of language development to understand and learn basic linguistic principles. For the linguistically more advanced child the robot offers the possibility of understanding basic properties of dialogue such as turn-taking, asking and answering questions, the importance of providing sufficient information, and cooperating to achieve a shared goal. In addition, the child learns to plan its actions in order to achieve a goal; e.g., getting the robot to draw a flower.

At yet more advanced stages, the child may learn simple "programming" to get the robot to repeatedly perform a complex action. For example, the child may provide a step-by-step instruction for drawing a square, and then name this shape *"square"*. Subsequently, the robot can be told to draw new squares using a single command (*"draw a square"*). This provides further practice in using dialogue to achieve more complex goals.

As discussed in section 3.2 later, the setup works without the robot and the communication board actually listening to each others' speech – instead, they communicate wirelessly. However, there is an important pedagogical point in having them (apparently) communicate using spoken language. It provides the child with an experience of participating in a spoken dialogue, even though the child does not speak.

2.3 Generality of the approach

One reason for choosing a drawing robot is that is provides a simple yet infinitely variable arena of behaviour. A further reason is that no advanced sensors or motors are needed to build such a robot. An alternative which is equally understandable and useful to the user could be a robot building towers using wooden blocks, but in this case the robot would need to be more advanced and difficult to construct.

This does not mean that the technique cannot be applied to other domains. There is nothing about the idea itself – a talking communication board communicating with a robot via a dialogue system – which dictates what the robot can be used for. To adapt the setup to a new domain, one needs to specify the relevant domain knowledge to the GoDiS dialogue system, and perhaps provide new signs for the communication board which are appropriate to the new domain.

3 Implementation

This section describes some technical aspects of the implementation of the TRIK system.

3.1 Components

The final TRIK setup consists of the following components:

- a simple LEGO robot which can turn and move in all directions, and has a pen that can be lifted and lowered;
- a touch-screen which functions as a communication board with pictograms/symbols;
- a computer with a dialogue system and speech synthesis, which is physically attached to the communication board and communicates wirelessly with the robot.

The computer will seem like it is a part of the communication board, but it also controls the robot, both movements and speech. Every utterance by the robot will be executed by the speech synthesizer, and then sent to the robot via radio.

3.2 Perfect speech recognition

Typically, the most error-prone component of a spoken dialogue system is speech recognition; i.e., the component responsible for correctly interpreting speech. This of course becomes even more problematic when working with language learning or communication disorders, since in these sitations it is both more difficult and more important that the computer correctly hears and understands the user's utterances. An advantage of the TRIK setup is that we will, in a sense, have "perfect speech recognition", since we are cheating a bit. The (dialogue system connected to the) robot does not actually have to listen for the speech generated by the (computer connected to the) communication board; since the information is already electronically encoded, it can instead be transferred wirelessly. This means that the robot will never hear *"go forward and then stop"* when the communication board actually says *"go forward seven steps"*.

3.3 Existing resources

This section describes the technical resources which are used in TRIK.

3.3.1 The GoDiS dialogue manager

A dialogue system typically consists of several components: speech recognizer, natural language interpreter, dialogue manager, language generator, speech synthesizer and a short-term memory for keeping track of the dialogue state. One can make a distinction between dialogue *systems*, which (ideally) are general and reusable over several domains, and dialogue system *applications*, which are specific to a certain domain. The dialogue manager is the "intelligence" of the system, keeping track of what has been said so far and deciding what should be said next.

The GoDiS dialogue manager (Larsson, 2002) is designed to be easily adaptable to new domains, but nevertheless be able to handle a variety of simpler or more complex dialogues. For example, GoDiS can either take initiative and prompt a user for information, or take a back seat and let the experienced user provide information in any desired order, without having to wait for the right question from the system.

3.3.2 The grammar formalism GF

Grammatical Framework (GF) (Ranta, 2004) makes it easy to quickly design the language interpretation and generation components of a dialogue system. In addition, GF is a multilingual formalism, which means that it is well suited for use in translation between different languages. Since, e.g., the graphical Blissymbolics system can be regarded as a language in itself, it is possible to write GF grammars for translating between symbols and spoken Swedish (Lidskog, 2007).

3.3.3 LEGO Mindstorms

The robot itself is built using LEGO Mindstorms NXT,[1] a kind of technical lego which can be controlled and programmed via a computer. Apart from being cheap, this technology makes it easy to build a prototype and to modify it during the course of the project.

4 Evaluation

During April–June 2009, the system will be evaluated by a number of users with linguistic communication disorders.

4.1 Design

The evalation process is designed as a case study with data being collected before and after interventions. The children will also be video recorded when playing with the robot, to enable analysis of common interaction patterns.

4.2 Users

The users will consist of children with a diagnose within the autism spectrum, and children with a CP diagnosis. The chronological age of the children may vary but the intention is to both include children in an early stage of language development, and children who have developed further and where there is a need to develop and train grammatical skills.

4.3 Evaluation method

After the children's families and/or personnel have been instructed about the use of the robot, they will be using it during 2 months. The children will have the opportunity to play with the robot about 2 to 3 times per week.

Before the robot is used, the parents answer a survey about how they perceive their interaction with their children. They will also estimate the communicative abilities of their children. The surveys will be complemented with questions based on the vocabulary which will be included in the children's communication boards. When the trial period is over, the surveys are repeated. During the trial period, the children will be filmed twice while using the robot, in the beginning and towards the end. The videos will then be analysed using suitable methods, such as Activity-Based Communication Analysis, developed at the University of Gothenburg. Furthermore, all interaction between the communication board and the robot will be logged by the system, providing valuable information to include in the overall analysis.

Acknowledgements

We are greatful to three anonymous referees for their comments. The TRIK project is financed by the Promobilia foundation and Magn. Bergvall's foundation.

References

Jonas Beskow, Olov Engwall, Björn Granström, and Preben Wik. 2004. Design strategies for a virtual language tutor. In *INTERSPEECH 2004*.

Mikael Heimann and Tomas Tjus. 1997. *Datorer och barn med autism*. Natur och Kultur.

Staffan Larsson. 2002. *Issue-based Dialogue Management*. Ph.D. thesis, Department of Linguistics, University of Gothenburg.

Johanna Lidskog. 2007. Swedish Bliss: Grammar based translation from Swedish into Bliss. Master's thesis, University of Gothenburg.

Seymour Papert. 1993. *Mindstorms: Children, Computers, and Powerful Ideas*. Basic Books.

Aarne Ranta. 2004. Grammatical Framework, a type-theoretical grammar formalism. *Journal of Functional Programming*, 14(2):145–189.

Stephanie Seneff, Chao Wang, and Julia Zhang. 2004. Spoken conversational interaction for language learning. In *InSTIL/ICALL 2004 Symposium on Computer Assisted Learning*.

Rose Sevcik and Mary Ann Romski. 2002. The role of language comprehension in establishing early augmented conversations. In I. J. Reichle, D. Beukelman, and J. Light, editors, *Exemplary Practices for Beginning Communicators*, pages 453–475. Paul H. Brookes Publishing.

Gunilla Thunberg. 2007. *Using speech-generating devices at home*. Ph.D. thesis, Department of Linguistics, University of Gothenburg.

[1] `http://mindstorms.lego.com/`

ISSN 1736-6305 Vol. 4
http://hdl.handle.net/10062/9206