"Artificial Communication"
Can Computer Generated Speech Improve Communication of Autistic Children?

Eberhard Grötsch[1], Alfredo Pina[2], Michael Schneider[1] and Benno Willoweit[1]
[1]*Fakultät für Informatik und Wirtschaftsinformatik, Hochschule für Angewandte Wissenschaften, Sanderheinrichsleitenweg 20, Würzburg, Germany*
[2]*Departamento de Informática, Universidad Pública de Navarra, Campus Arrosadia, Pamplona, Spain*

Keywords: Natural Language Processing, Autistic, Autonomous Toy, Robot, Children.

Abstract: Autistic children are often motivated in their communication behaviour by pets or toys. Our aim is to investigate, how communication with "intelligent" systems affects the interaction of children with untypical development. Natural language processing is intended to be used in toys to talk to children. This challenging Háblame-project (as part of the EU-funded Gaviota project) is just starting. We will discuss verification of its premises and its potentials, and outline the technical solution.

1 INTRODUCTION

It is a well established fact that autistic children often are motivated in their communication behaviour by pets or toys, e.g. in the IROMEC project (Ferari, Robins, Dautenhahn, 2009), (IROMEC, 2013). We found analogous results in a group of disabled persons who were motivated by technical systems to move or dance. (Pina, 2011).

Within the Gaviota Project (Gaviota, 2012), we want to investigate, how communication with "intelligent" systems affects the interaction of children with untypical development.

2 PREVIOUS WORK

2.1 The Beginning: Eliza

As early as 1966 Weizenbaum (Weizenbaum, 1966) implemented an interaction technique which was introduced by Carl Rogers (client centered psychotherapy, (Rogers, 1951)). This therapy mainly paraphrases the statement of the client. The Eliza implementation used to react to a limited number of key words (family, mother, ...) to continue a dialog. Eliza had no (deep) knowledge about domains - not even shallow reasoning, rather a tricky substitution of strings. Modern versions of Eliza can be tested on several websites, e.g. (ELIZA, 2013).

2.2 Robots in Autism Therapy

So far robots in autism therapy have been used to enhance the abilities of children to play, using robots as a toy, which means they playfully interact with robots.

The robot's simple face can be changed to show feelings of sadness or happiness by different shapes of the mouth (IROMEC, 2013).

These robots (which are just special computer screens in a first step) execute pre-defined scenarios of interaction, and are controlled by humans.

So far results have shown that more children are responding to those robots compared to the children that do not respond.

2.3 State-of-the-Art Dialog Systems

State of the art dialog systems (e.g. the original Deutsche Bahn system giving information about train time tables, or the extended system by Philips) are able to guide people who call a hotline and execute standardized business processes (delivering account data, changing address data, etc.). Those systems work well, but within an extremely limited domain.

Grötsch E., Pina A., Schneider M. and Willoweit B..
"Artificial Communication" - Can Computer Generated Speech Improve Communication of Autistic Children?.
DOI: 10.5220/0004412805170521
In *Proceedings of the 5th International Conference on Computer Supported Education* (CSEDU-2013), pages 517-521
ISBN: 978-989-8565-53-2

2.4 Natural Language Processing (NLP)

A spectacular demonstration of natural language processing was given by IBM's artificial intelligence computer system Watson in 2011, when it competed on the quiz show Jeopardy! against former human winners of that popular US television show (JEOPARDY, 2011).

IBM used the Apache UIMA framework, a standard widely used in artificial intelligence (UIMA, 2013). UIMA means "Unstructured Information Management Architecture".

UIMA can be viewed from different points of view:

1) architectural: UIMA represents a pipeline of subsequent components which follow each other in an analytical process, to build up structured knowledge out of unstructured data. UIMA primarily does not standardize the components, but the interfaces between components.
   "*... for example "language identification" => "language specific segmentation" => "sentence boundary detection" => "entity detection (person/place names etc.)". Each component implements interfaces defined by the framework and provides self-describing metadata via XML descriptor files. The framework manages these components and the data flow between them. Components are written in Java or C++; the data that flows between components is designed for efficient mapping between these languages*". (UIMA, 2013).
2) UIMA supports the software architect by a set of design patterns.
3) UIMA contains two different ways of representing data: a fast in-memory representation of annotations (high-performance analytics) and an XML representation (integration with remote web services).

The source code for a reference implementation of this framework is available on the website of the Apache Software Foundation.

Systems that are used in medical environments to analyze clinical notes serve as examples.

2.5 Natural Language Processing in Pedagogics

So far there are no reasoning systems with knowledge about the domain of how to behave properly in a pedagogical way.

3 HYPOTHESIS: NATURAL LANGUAGE SPEAKING MIGHT BE HELPFUL

The IROMEC project demonstrated that weekly sessions with a robot with rather simple abilities to move and show emotions by standardized facial expressions are helpful to enable/empower children to play more naturally than without those sessions (Ferari, Robins, Dautenhahn, 2009). So we concluded that it is worth trying to build a robot, which is talking autonomously with a child in rather simple and standardized words and sentences. We decided to start a subproject Háblame („talk to me") to investigate the chances and problems of building such a robot as part of the EU-funded Gaviota project.

4 THE PROJECT „HÁBLAME"

4.1 Verification of the Hypothesis

Before we start the core project, we have to verify our hypothesis: we have to show that autistic children positively react to toys which talk to them. We will build a simple prototype without NLP-functions. Speech will be produced by a hidden person via microphone and suitably placed speakers.

4.2 Concept of a Dialog System

Within the project, we first had to / have to get experience with natural language processing. When we studied basic concepts of NLP (Figure 1), we decided to put stress on syntax parsing and semantic parsing.

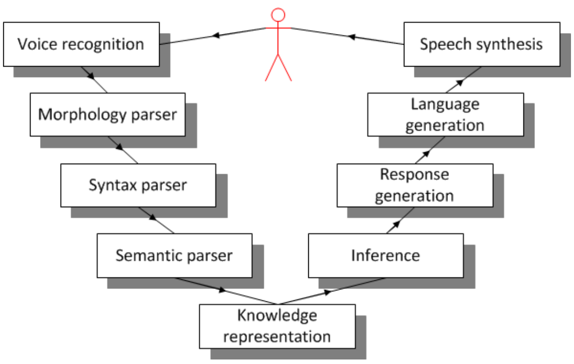

Figure 1: Concept of a dialog system (Schneider, 2012).

4.3 Parsing Syntax of Natural Languages

First a prototype parser – based on a grammar developed by Roland Hausser (Hausser, 2000) – was implemented, which can analyze simple sentences entered in English. The parser processes the sentence entered, splits it into words and compares them to a lexicon, specified in an external text file.

It tries to re-combine the sentence word by word, taking account of the *valences*, also specified in the lexicon. If the sentence can be re-combined correctly and all free valences are filled, the parsing process was successful. Otherwise the sentence is grammatically incorrect (or the parser could not deal with it).

4.3.1 Parser Prototype and Valences

The parser works with valences of words, e.g.:

- to sleep has 1 nominative valence → *Peter sleeps.*
- to give has 1 nominative valence (abbreviated Nx), 1 dative valence (Dx) and 1 accusative valence (Ax) → *Peter gives Mary books.*
- All valences (mostly opened by verbs) have to be filled (mostly by nouns). Otherwise the sentence is not correct, e.g.: *Peter gives Mary.* → accusative noun is missing.

One can think of valences as slots, which have to be filled with proper words.

4.3.2 Processing Valences

Valid words, their valences and their function (V = verb, PN = plural noun, etc.) have to be specified in an external lexicon, e.g.:

- *sleeps NS3x V*
  (S3: use only with 3rd person singular)
- *give N-S3x Dx Ax V*
  (-S3: use NOT with 3rd person singular)
- *books PN*

Words currently have to be entered in the lexicon with all flection forms used, e.g.:

- *give N-S3x Dx Ax V*
- *gives NS3x Dx Ax V*
- *gave Nx Dx Ax V*

The parser takes the first word of the sentence and combines it with the following word to a more complex starting sequence using predefined rules, e.g.:

- Noun phrase followed by a verb with corresponding valence → erase the valence satisfied:
  Peter (SNP) sleeps (NS3x V). →
  Peter sleeps (V).
- Article followed by adjective → do not change any valences:
  The (SNx SNP) beautiful (ADJ) … →
  The beautiful (SNx SNP) …

This combining procedure is repeated bottom-up until the end of the sentence is reached (Figure 2).

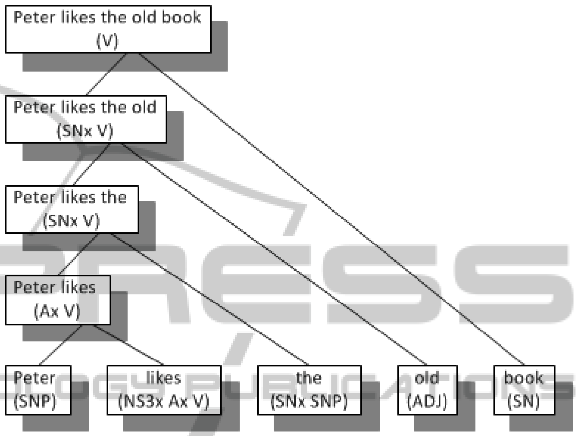

Figure 2: Bottom-up processing of valences (Schneider, 2012), cf. (Hausser, 2000).

Examples of sentences, the prototype of the parser can deal with:

- *The beautiful girl reads an old book.*
- *Does Peter sleep?*
- *Mary has bought a new car.*

Examples of sentences, the prototype currently cannot deal with:

- *Beautiful girls like Peter.*
- *Reading books gives Peter pleasure.*
- *Peter, who is 20 years old, sleeps.*

4.4 Processing of Semantics of Natural Languages – Analyzing Semantics

Analyzing the semantics of natural language, we first define our prerequisites and our goals:

- Prerequisites:
  Oral utterances (of children) are transcribed by a supervisor and fed into the system. The sentences are analyzed one by one, and the results of the analysis should be stored in a semantic network
- Goals:
  - Exploring the linguistic techniques for semantic analysis.
  - Determining the technical and linguistic preconditions.

- Evaluate which software components and libraries may be used to accomplish this task.
- Evaluate which libraries can be used to access a semantic network, and how to create the necessary ontologies.
- Building a software prototype, which integrates all necessary components.

Basically, there are two approaches towards linguistic analyzing:

The „**formal**“ approach:

Every sentence represents a logical statement („Proposition“), and we have to translate every sentence into meta-language. Those languages are called „Meaning Representation Languages“ (MRL) and are often based on first order logic or the lambda calculus.

The „cognitive“ approach:

One can‘t determine the exact meaning of a sentence by the sentence itself. A straightforward translation of language into a logical representation is therefore impossible.

In the process of understanding there is a lot of background knowledge involved.

This knowledge may be specific to a single person or a group of persons (e.g. cultural or personal background).

4.4.1 Adoption in Computational Linguistics

The *formal* approach is well explored and adopted in Computational Linguistics.

Its main advantages are easy integration with code and other logical structures like semantic networks. The disadvantage is that it is not language agnostic and very narrow in scope (one has to define logical expressions for every meaning of a sentence).

The cognitive approach was investigated mainly by adopting Fillmore‘s work on frame semantics, which he developed back in the 1970s (Fillmore, 2006). His idea was that the meaning of a sentence can be described by a so-called frame or a combination of those. A frame is consisting of:

- A description which outlines the meaning of the frame
- A number of frame elements (FE) that describe possible roles or agents
- Relations to other frames, including specialization, part-of or temporal relations
- A number of language specific lexical units, i.e. words or groups of words, which may evoke that frame.

The main advantage of the cognitive, frame-based approach is, that frames are language agnostic, so only the lexical units that may evoke a frame have to be defined per language. Every frame is a formal representation of meaning, so there is no reason to build an own meta-language. The scope is very broad and not limited to a specific application.

4.4.2 Software Tools for FrameNet based Analysis (Cognitive Approach)

The FrameNet database consists of a large set of XML files (FrameNet, 2012).

Frame semantic parsers relying on FrameNet already exist, both systems use a probabilistic approach:

- SHALMANESER (English, German) is a project at Saarland University, Saarbrücken, Germany, and
- SEMAFOR (English) is a project at Carnegie Mellon University, Pittsburgh, USA.

4.4.3 Preprocessing of Sentences (Cognitive Approach)

In a first step we preprocess the sentences to be analyzed:

- Tokenizing: we split sentences into words (Apache NLP Tools),
- POS-Tagging: we determine the part of speech of each token (Apache NLP Tools),
- Syntactic parsing: Determining the grammatical components of each sentence (Maximum Spanning Tree Parser, Pennsylvania State University),
- Named Entity Recognition: Check if one or more tokens represent a proper noun, a number, a date, etc. (Apache NLP Tools),
- Frame identifications: Find the frames that match the given sentence (Semafor, Carnegie Mellon University, Pittsburgh, USA).

5 RESULTS

So far there are only results, as far as NLP is concerned:

- The pre-trained classifiers for both SHALMANESER and SEMAFOR did not yield good results with our test data.
  - SHALMANESER is hard to integrate with other tools.
  - There are plenty of java-based tools to preprocess the data and extract features that can be used with probabilistic models. Fur-

thermore, many of these tools can be integrated with the Apache UIMA platform.
- A modular, client/server based approach proved to be necessary for the project.
- A fairly large corpus of transcribed child language is nearly impossible to obtain.
- Although there are FrameNet data sets for a couple of languages (Spanish, German, Chinese, etc.), their number of frames and lexical units is presumably too small to use for semantic parsing.

6 CONCLUSIONS

First we have to verify that autistic children react to the prototype system in the manner expected.
If this is done successfully, there is much work left to be done on the NLP side. We will not do further research on using FrameNet with the Semafor parser however, nor use *database semantics* (another approach, which is not covered in this report).
We will intensify research on custom probabilistic models with the following steps:

1. set up Apache UIMA since the NLP tools are easy to integrate,
2. obtain a domain specific corpus,
3. split that corpus into a training and a test part,
4. annotate the corpus with semantic class labels,
5. select domain specific and situational features,
6. incorporate the features generated by the preprocessing tools (i.e. taggers, parsers, etc.),
7. train a probabilistic model, possibly by using the MaxEnt library of the Apache NLP tools,
8. evaluate the performance with different feature sets.

6.1 Necessary Data

We need corpora about children's language domains, and we have to decide, which age level, and which speech domains. If no corpus is available, we have to develop one. Those corpora should be in English language to develop and stabilize the system. Later iterations may incorporate German and Spanish language.

6.2 Further Steps

We will set up an experimental environment, based on the work already done, gather experience and knowledge on analyzing/parsing natural language. Then we have to acquire or produce corpora covering our domain of interest (child language). Furthermore we have to work on creating natural sentences as part of a dialog.

ACKNOWLEDGEMENTS

This work has been partially funded by the EU Project GAVIOTA (DCI-ALA/19.09.01/10/21526/245-654/ALFA 111(2010)149.

REFERENCES

ELIZA, 2013. www.med-ai.com/models/eliza.html (March 3,2013)

Gaviota, 2012. Report on the Results of the Gaviota Project, International Meeting, Santa Cruz, Bolivia (unpublished presentations)

Ferari, E., Robins, B., Dautenhahn, K., 2009. Robot as a Social Mediator - a Play Scenario Implementation with Children with Autism, *8th International Conference on Interaction Design and Children Workshop on Creative Interactive Play for Disabled Children*, Como, Italy

Fillmore, C. J., 2006. Frame Semantics, in Geeraerts, D. (ed.): Cognitive Linguistics - Basic Readings, chap. 10, Mouton de Gruyter, p. 373–400.

FrameNet, 2012. The FrameNet Project, University of California, Berkeley, https://framenet.icsi.berkeley.edu (Mar 06, 2013)

Hausser, R., 2000. Grundlagen der Computerlinguistik – Mensch-Maschine-Kommunikation in natürlicher Sprache, Springer Verlag Berlin

IROMEC, 2013. http://www.iromec.org/9.0.html (Jan 27, 2013)

JEOPARDY, 2011. http://www.nytimes.com/2011/02/17/science/17jeopardy-watson.html?_r=0, (Jan 28, 2013)

Pina, A., 2011. New Technologies for Language and Learning Disabilities, *17th International Conference on Technology supported Learning & Training*, Online Educa Berlin

Rogers, C. R., 1951. Client-centered therapy, Oxford, Houghton Mifflin

Schneider, M., 2012. Processing of Semantics of Natural Languages – Parsing Syntax of Natural Languages, Bachelor-Thesis, HAW Würzburg-Schweinfurt

UIMA, 2013. http://uima.apache.org/ (Jan 28, 13)

Weizenbaum, J., 1966. ELIZA - A Computer Program for the Study of Natural Language, Communication between Man and Machine. Communications of the ACM. New York 9.1966,1. ISSN 0001-0782

Willoweit, B., 2012. Processing of Semantics of Natural Languages – Analyzing Semantics, Bachelor-Thesis, HAW Würzburg-Schweinfurt