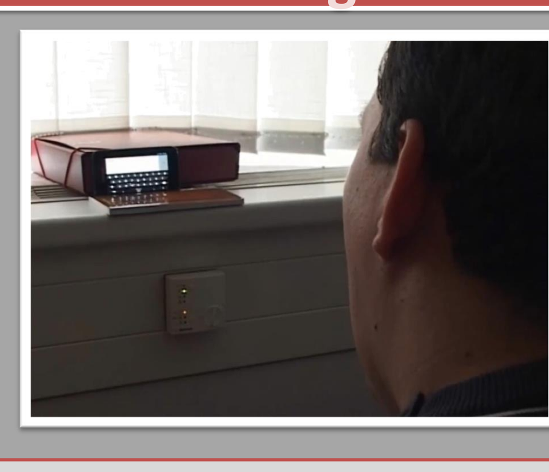
Gaze tracking and Language Model for Flexible Augmentative and Alternative Communication in Practical Scenarios

Eötvös Loránd University – and – German Research Center for Artificial Intelligence



MOTIVATION

(i) Symbol based communication: continuous confirmation of **understanding** is needed
(ii) Possibilities of **initiating conversations** are restricted
(iii) We **cannot assume a grammatical structure** of symbols
(iv) Communication with most people is difficult.

CHALLENGES

AAC symbol boards are slow and cumbersome,
General situations are problematic
It is hard to make personal contacts.
Symbol (word tag) combinations are ambigous.

CONTRIBUTION

(i) We generate natural language sentences from selected word (symbol tag) sequences.
(ii) Symbol selection is facilitated by gaze tracking.

ALGORITHM DETAILS

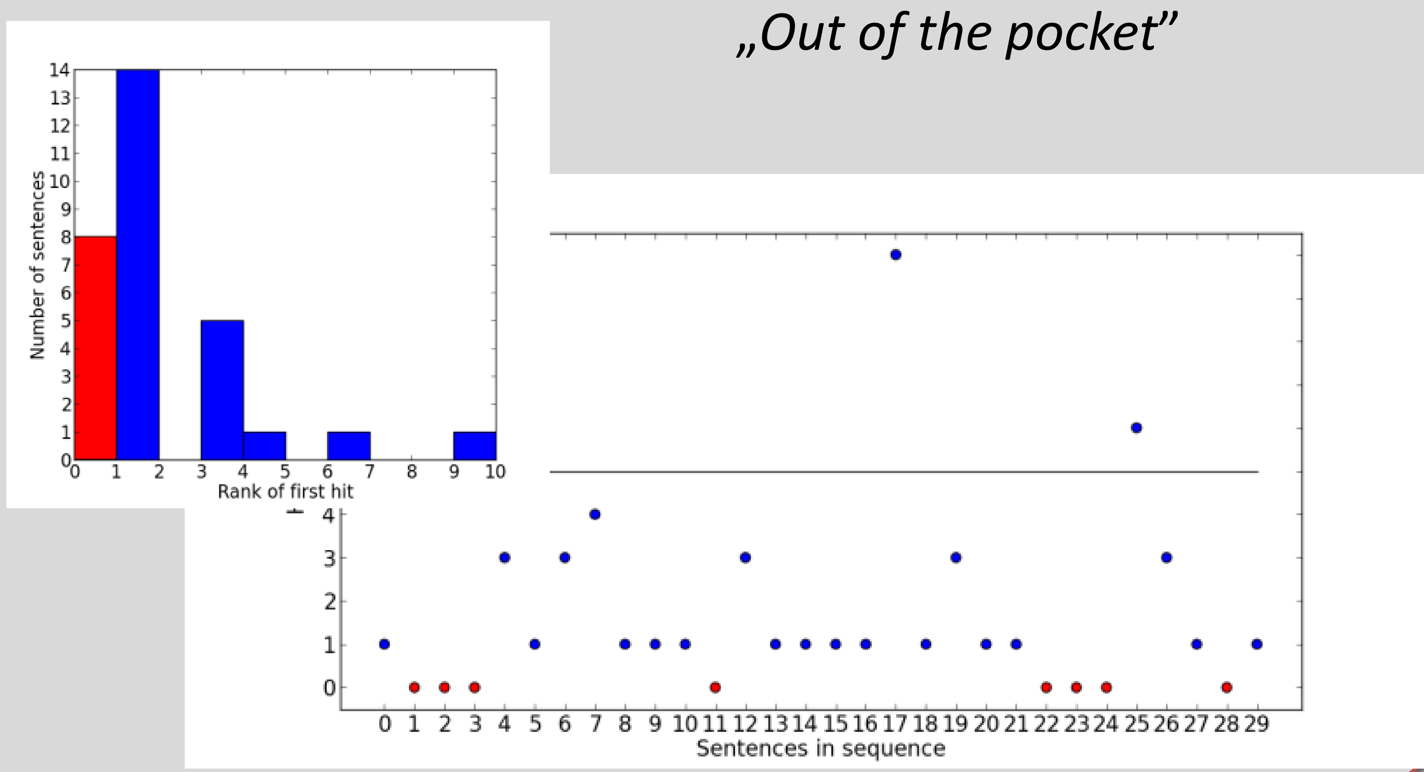
(i) We use **Statistical Language Modeling** to generate full sentences.
(ii) We generate sequences consisting of **symbols** and **closed words**.
(iii) We estimate the probability of a sequence using an **N-gram** language model.
(iv) Since the number of the possible combinations of symbols and closed words tends to increase exponentially therefore we developed a greedy algorithm to restrict the candidate space, that we call **Suffix Tree Builder**.
(v) For estimating language model parameters we use the huge and general **Google Books N-gram Corpus** with the **BerkeleyLM Language Model Library**.

FUTURE WORK

(i) Incorporate context based filtering.
(ii) Multimodal information extraction methods.
(iii) Gther semantic information of the situation: process the additional video and audio data.

RESULTS

Algorithm: rank syntactically correct sentences for each word sequence during three scenarios. Scenarios include thirty word sequences. Blue: hit, Red: no hit
73% of the results contained appropriate sentences
91% of the correct sentences were in the first five hits.
Example: {„*out*", „*pocket*} → „*Out of my pocket*" or „*Out of the pocket*"



ACKNOWLEDGEMENTS

Grant: Research and Technology Innovation Fund. EITKIC 12,

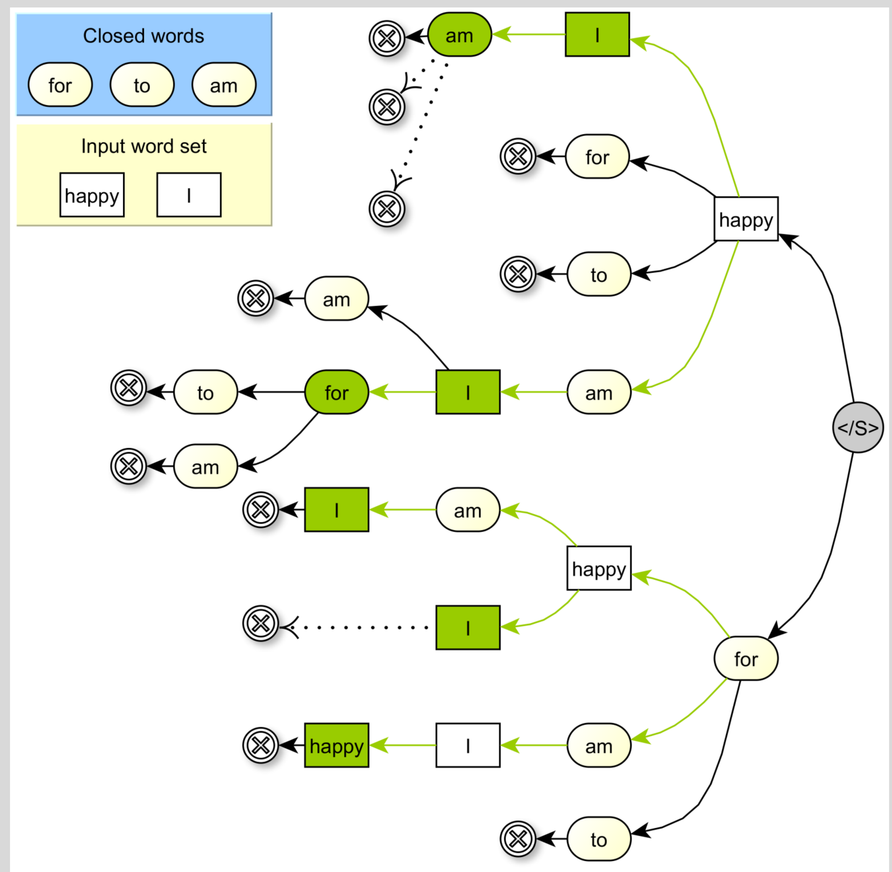
Anita Verő*, Brigitta Miksztai-Réthey, Gyula Vörös, Ádám Zsigmond, Balázs Pintér, Takumi Toyama, Jason Orlosky, Daniel Sonntag, András Lőrincz**
*anitaveroe@gmail.com, **lorincz@inf.elte.hu

German Research Center for Artificial Intelligence GmbH
EIT ICT Labs

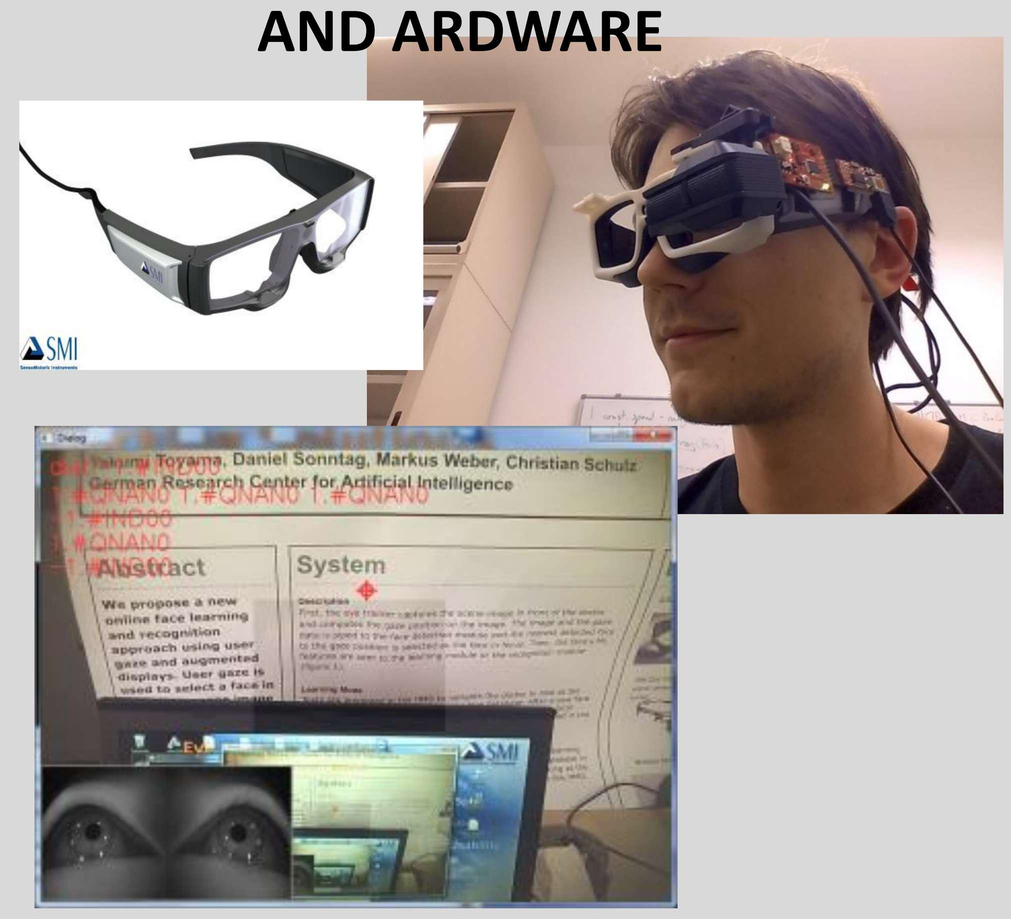
METHODS

Traditional: head stick or headmouse
More natural tools:
Eye Tracking Glasses (SMI ETG) : a pair of glasses with eye tracking infrared cameras and a forward looking video camera.
Head Mounted Display (Brother HMD) : see-through display attached to glasses.

SUFFIX TREE BUILDER ALGORITHM

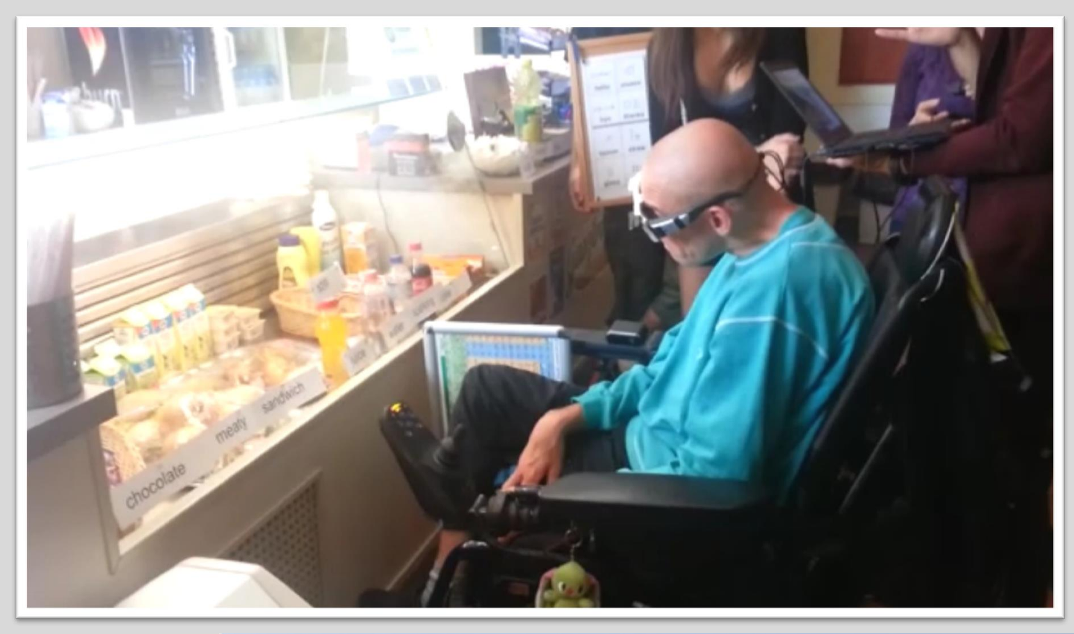


EYE TRACKING SOFTWARE AND ARDWARE

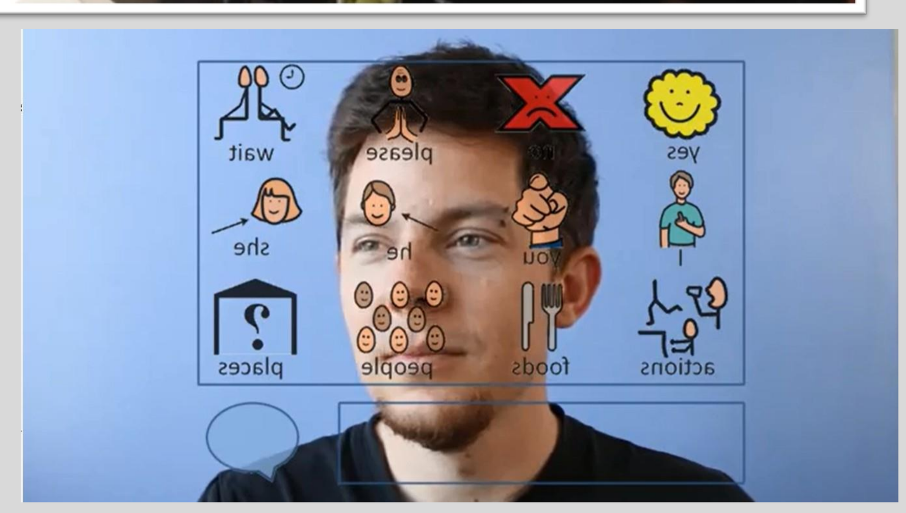


The algorithm builds the sentence fragments step-by-step by traversing a pruned suffix tree of word sequences (i.e., n-grams). Utilizes probabilities.
Nodes with **rounded shape** correspond to **closed words, green color** indicate the first word of an accepted fragment.

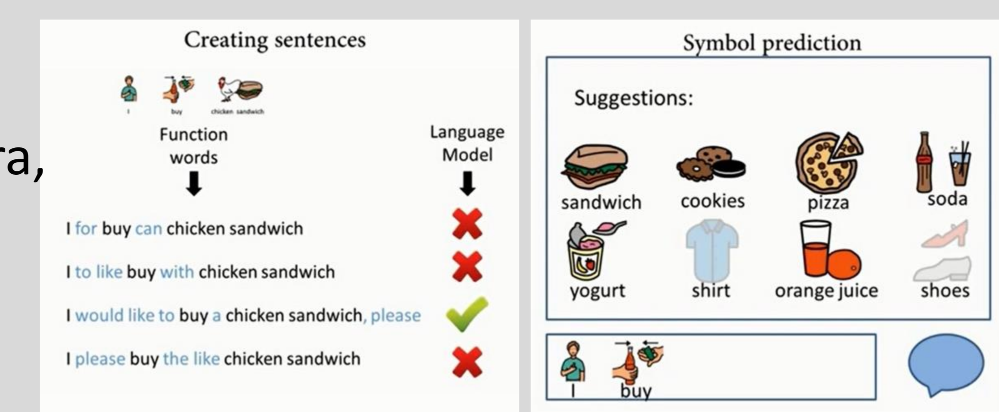
PRACTICAL EXPERIMENTS



TESTS IN A FOOD STORE:



Communication board is printed on a piece of paper.
Optical character recognition is used to let the participant combine the symbols on the board and symbols in the external world.
System tracks eye movements.
Glasses contain a forward looking camera, and two eye monitoring cameras.



AAC TELEPRESENCE:

Head controlled **Android Based Robot** using **IvenSense MPU-9150 gyroscope,** and see-through **Moverio BT-100 head-mounted display** (HMD): user can (a) look at the partner, (b) select symbols, (c) generate sentences, (d) tell! them through the mobile.