Recommending Missing Symbols of Augmentative and Alternative Communication by Means of Explicit Semantic Analysis

Gyula Vörös, Péter Rabi, Balázs Pintér, András Sárkány
Faculty of Informatics, Eötvös Loránd University
H-1117 Budapest, Hungary

Daniel Sonntag
German Research Center for Artificial Intelligence
66123 Saarbruecken, Germany

András Lőrincz
Faculty of Informatics, Eötvös Loránd University
H-1117 Budapest, Hungary

Abstract

For people constrained to picture based communication, the expression of interest in a question answering (QA) or information retrieval (IR) scenario is highly limited. Traditionally, alternative and augmentative communication (AAC) methods (such as gestures and communication boards) are utilised. But only few systems allow users to produce whole utterances or sentences that consist of multiple words; work to generate them automatically is a promising direction in the big data context. In this paper, we provide a dedicated access method for the open-domain QA and IR context. We propose a method for the user to search for additional symbols to be added to the communication board in real-time while using access to big data sources and context based filtering when the desired symbol is missing. The user can select a symbol that is associated with the desired concept and the system searches for images on the Internet—here, in Wikipedia—with the purpose of retrieving an appropriate symbol or picture. Querying for candidates is performed by estimating semantic relatedness between text fragments using explicit semantic analysis (ESA).

Introduction

The ability to communicate with others is of paramount importance for mental well-being. People with severe speech and physical impairments (SSPI) such as cerebral palsy, stroke, ALS/motor neurone disease (MND), or muscle spasticity face enormous challenges in daily life activities. A person who cannot speak may be forced to only communicate directly to his closest relatives, thereby completely relying on them for interacting with the external world.

Some systems allow users to produce whole utterances or sentences that consist of multiple words. The main task of the AAC system is to store and retrieve such utterances. However, using a predefined set of sentences severely restricts the applicability: in many cases, such as picture based communication, the expressivity is rather limited because for practical reasons, communication boards (see figure 1) can only contain a small set of symbols. Utterances consisting of multiple symbols are often telegraphic: they are unlike natural sentences, often missing words to allow for successful and detailed communication, "Water now!" instead of "I would appreciate a glass of water, immediately."

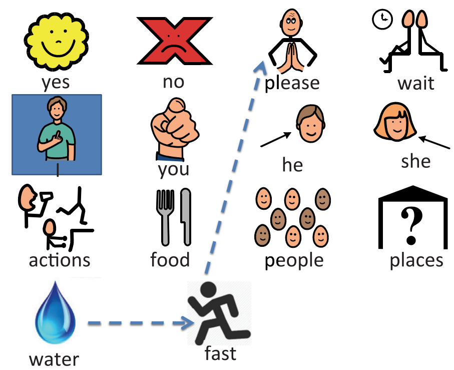


Figure 1: Communication board example with 14 symbols.

If the symbol set is large, selection becomes slow. On the other hand, the symbol set should be as specific as possible to allow for open-domain communication required when addressing open-domain data sources for information retrieval and question answering.

Based on a wearable communication board system (Verő et al. 2014; Vörös et al. 2014) made of (i) the head-mounted human-computer interaction part consisting of smart glasses with gaze trackers and text-to-speech functionality, which implement a communication board and the selection tool, and (ii) a natural language processing pipeline in the backend in order to generate complete sentences from the symbols on the board, in this paper, we report on a new approach of recommending missing symbols of AAC by means of explicit semantic analysis.

In AAC, the number of symbols available to a person at a time is very limited, in particular in the open-domain QA and retrieval context: often a symbol for the desired open-domain concept does not exist at all, or it is hard to find the symbol in the very large set of symbols. (Or there is no symbol for the specific concept since it describes a very special personal experience that would not generalize to communication board symbols.) We propose a method for the user to search for additional symbols or pictures in real time using big data and context based filtering when the desired symbol is missing on the communication board.

In this paper, we take a step in the direction of easing AAC communication. We want to give the user opportunity to automatically expand his or her communication board: when a symbol for a desired and given linguistic concept (the *cue*) is missing, a context-dependent additional symbol set should be made available in an ad hoc fashion to exceed the limit of the initial communication board, based one available online resources. In our approach, the user can select a symbol (the *target*) that is associated with the cue concept and the system searches for images on Wikipedia with the purpose of including an image to the communication board corresponding to the cue. We assume the user employs *free association* (Nelson, McEvoy, and Dennis 2000) to select a target symbol for the cue concept. In our experiments, we used the USF Free Association Norms (Nelson, McEvoy, and Schreiber 1998) to simulate a user action looking for a symbol. A key issue is to estimate the reverse free association strength (Nelson, Dyrdal, and Goodmon 2005). We use Explicit Semantic Analysis (ESA) (Gabrilovich and Markovitch 2007; 2009) to estimate semantic relatedness of the retrieved images to the target symbol. Our first experiment examines the possibilities in the method by showing how effectively can free association be estimated with ESA.

If the symbol for the cue concept does not exist in the set, it can be necessary to search for appropriate images from other sources. In our second experiment, we tested the system with the first images from Wikipedia pages. In this experiment, human judges evaluated whether our system returned appropriate images or not.

Background and Related Work

Symbol Based Communication

The set of symbols available to the user (i.e., the vocabulary) can usually be accessed using a hierarchical communication board with categories and hyperlinks. These structures enable users to access a large number of symbols, but as the vocabulary becomes larger, it is increasingly harder to traverse. Even if the required symbol is present in the system, it may be hidden in a deep hierarchy, and the user might forget its location. Another related problem is that symbol sets themselves are limited. Language is constantly changing and expanding, so there is a constant need to include new words which do not exist on a static communication board. Symbols for personal experiences or photos can also be added to the communication board. In conclusion, the limited number of symbols that are made available and accessible to the user at given time point (context situation) limits the expressivity of board communication.

A widely used technique to overcome these difficulties is *rate enhancement*, i.e., to enable the user to communicate faster. One method is to predict the next symbol of the currently edited message (Wiegand and Patel 2012). The vocabulary structure can also be enriched with suggestions (Nikolova, Tremaine, and Cook 2010). The context (e.g., GPS coordinates) can also be used to retrieve a vocabulary for the given situation. These context-based vocabularies are usually manually crafted, (Kim et al. 2009; Epp et al. 2011); however, some work has been done to generate them automatically (Demmans et al. 2012).

High-tech tools which facilitate symbol-based communication may provide means to assemble messages from multiple symbols, and generate natural language utterances from the messages. There are many methods that take advantage of natural language processing, take for example, Niki Talk, Avaz AAC App, Proloquo2Go, or Tobii Sono Flex, among many other innovative developments. As introduced, we are particularly interested in the combination of natural language processing methods and context based filtering by using high-tech smart mobile tools (Verő et al. 2014; Vörös et al. 2014) for natural language access within multimodal dialogue systems (Sonntag et al. 2010; 2007).

Approach

Consider a specific communication setting: a person who is using a communication board cannot find a symbol for expressing a specific information need. In the example in figure 2 (a), the context is a dining room, and the missing concept on the communication board is a knife. The concept ”knife” becomes the desired *cue* in our symbol selection method and the procedure for adding it to the symbol selection board is as follows:

1. Select another “related” symbol that we call *target* by free association (related means whatever symbol in the list of available symbols comes into the user’s mind).
2. Make an informed guess of the desired *cue* and provide a visual representation of it. “Fork” is the only available symbol on the board related to eating. ESA provides many concepts as associations that includes “knife”, but many of them are irrelevant, such as the river that may have “forks” or “Lake Fork”, among others (figure 2 (b)). The dining room context can restrict the symbol set (figure 2 (c)).
3. If successful, then the original communication goal can be fulfilled.

Preparation

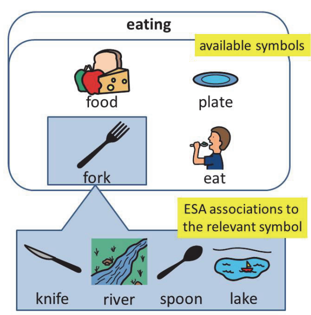
The method uses Wikipedia[1] as its knowledge base to extract both concepts as well as corresponding images, and we define a semantic similarity metric between pairs of expressions. We selected Wikipedia because of its comprehensive coverage of common-sense concepts and knowledge useful for communication board situations and because all images are free for non-commercial use. We used the XML dump[2], containing the MediaWiki formatted content of all pages.

First, we determine the most similar possible cues to the target given by the communicating person, ordered by the semantic similarity metric in descending order of similarity. The semantic similarity of expression pairs was determined by cosine similarity on ESA representations of the expressions which depict a certain concept to be visualized

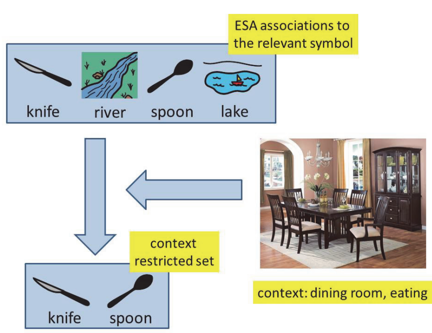
[1] http://en.wikipedia.org/wiki/Main_Page

[2] http://dumps.wikimedia.org/enwiki, we used the July 8, 2014 dump for our image collection process.

(a)



(b)



(c)

Figure 2: Context based filtering approach along the formulation of the selection task: the problem is that the concept “knife” is missing on the board.

according to (Gabrilovich and Markovitch 2007). ESA assumes that every page of Wikipedia is a concept and also that if a word is mentioned on a page, then it is related to the concept the page represents. The frequency of the word on the page shows the strength of this relation. For each word, an ESA *interpretation vector* is built by assigning a weight to each Wikipedia concept based on the frequency of the word on the corresponding Wikipedia page. The interpretation vector of a multi-word expression is the centroid of the vectors representing the individual words. Semantic relatedness between words and expressions is computed simply by taking the cosine similarity between their ESA interpretation vectors. For the details of the incorporated algorithm, see (Gabrilovich and Markovitch 2007; 2009).

Second, we present the corresponding images of the most similar Wikipedia concepts to the user: the first 24 most similar concepts in this list are presented (this is approx. the number of images that can be dealt with concurrently on a communicate board). The images representing the potential *cues* are obtained by downloading the first image from the page of each expression on Wikipedia. The user then can select an image from this set according to his or her query or intent.

Evaluation

We conducted two experiments to evaluate the approach. In these experiments, we first modelled the potential interesting associations for the communication board user with a dataset on direct human associations, the Free Association Norms dataset (Nelson, McEvoy, and Schreiber 1998) where participants were asked to write the first thing that came to mind that was meaningfully related or strongly associated to the presented word or multi-word expression on the blank shown next to each item. For example, if given $[BOOK, ...]$, they might write $READ$ on the blank next to it.

The set of concepts we work with on the communication boards have been selected based on the Picture Communication Symbols (PCS)[3] from BoardMaker v5. We used the associated word or expression belonging to each symbol as the concepts. When a symbol came with graphical variations, only one variant was included together with its associated concept set. PCS names starting with a digit, and symbols representing special signs such as letters were filtered out.

For the evaluation of our AAC recommendation approach we selected 2476 PCS symbols. We automatically filtered 572 associations from the Free Association Norms dataset that all satisfied the following criteria:

1. both cue and $target$ are in the PCS set
2. the individual cue of a target has got an image representation on Wikipedia.[4]
3. the cue has the *strongest connection* to the target[5]

[3] http://www.mayer-johnson.com

[4] The images were extracted as follows: (1) for each cue, the Wikipedia article about that cue word or expression was looked up by matching the titles of the articles with the names of the symbols; (2) matching was case-insensitive and whitespaces were replaced with underscores; (3) when we found a match, then the first image in the article—located in the infobox or in the text body—was downloaded as the image representation of the given cue.

[5] We used forward association relative frequency (FSG) as our measure.

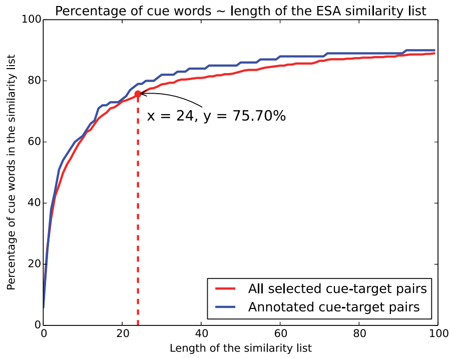


Figure 3: Results - Automatic testing.

First experiment

In the first experiment, we measured how suitable the candidate cues or expressions our system produces are (independently of the images assigned to them.) This experiment involves no supervised annotations and is conducted fully automatically.

Given a list of *cue*, *target* pairs that we collected according to the method detailed above, we composed the list of *potential cues* for each target of each pair based on ESA similarity, as described in the previous section. Specifically, one such list contains all of the words or expressions in the set of possible cues, ordered by their ESA similarities to the target in descending order.

We computed the rank of the correct cues in this list for each target. As the images for the communication board are chosen in this same order (i.e., according to their ESA similarity to the target), the ranks represent the ordinal number of the image of the correct cue in the selection process for the communication board. For example, a rank of 5 means that the correct image would be the fifth alternative on the communication board. A communication board that consists of 4 images would not contain the correct image, whereas a communication board that consists of 5 or more image would include it (consider again that we assume a board of approx. 24 images.)

The goal of this first test was to provide an intuition about the recommendation accuracies that can be obtained in optimal conditions: we assume that all the images can be easily extracted from the Wikipedia page, then the user could choose a suitable image for every cue for which an image is present on the communication board. With this assumption, we can compute the percentage of words or expressions for which the communication board would contain the correct image for a given communication board size (see the red line for the communication board with 24 symbols in figure 3).

We implemented the actual communication board based on the new selection approach and chose 24 symbols to be added to the experimental communication board for two reasons. First, as we found in this first experiment, the correct word or expression was present in the first 24 candidates 75% of the time; Second, as already mentioned 24 symbols seem to be a reasonable choice since it can fit on a board in 4 rows and 6 columns and visual search and selection can still be fast.

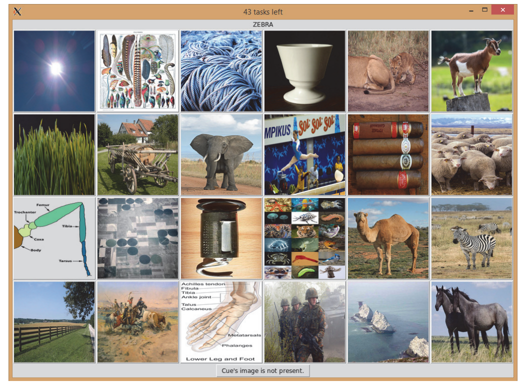


Figure 4: Desktop annotation tool for Wikipedia symbol suggestions. Cue: zebra, target: horse. If the cue is not present, then the annotator is supposed to press the button under the images.

Second experiment

In the second experiment, 6 independent human annotators were tasked to use the complete system on 100 examples.

A desktop feedback tool (figure 4) has been implemented and used by 6 independent annotators to “communicate” 100 randomly chosen cues out of the 572 filtered instances. The task was as follows: for each presented cue, they had to select one of the 24 presented images as being correct for this cue, if they found one. The words or expressions were the same for each annotator, in order to (i) produce aggregate statistics and (ii) measure inter-annotator agreement. The order in which the images were presented on the screen was random, but each annotator saw the same arrangement.

Inter-annotator agreement between the six annotators (Davies and Fleiss 1982) was $\kappa = 0.70$. This is a fairly high value for the 25 alternative forced choice (with one alternative being the “image is not present” answer). On average, the annotators found a suitable image in 64 cases out of 100, with a standard error of 1.46. They selected the image with the exact same Wikipedia label as the cue in 52 cases out of 100, with a standard error of 1.65. Figure 5 shows the results for the individual annotators.

Discussion

Word/concept-image correspondence

In the following, we discuss some special cases of the extracted Wikipedia images as symbol. Apparently they fall into three categories of different suitability of the word/concept-image correspondence:

First, figure 6 shows simple correspndences every annotator has indicated that an appropriate image has been found. It can be said that for example there is a high concept-image correspondence for animals and plants: all annotators agreed on appropriate image instances for *bee, fish, zebra, cat, gorilla, mouse, flower* and *tree* for example. Prototype images

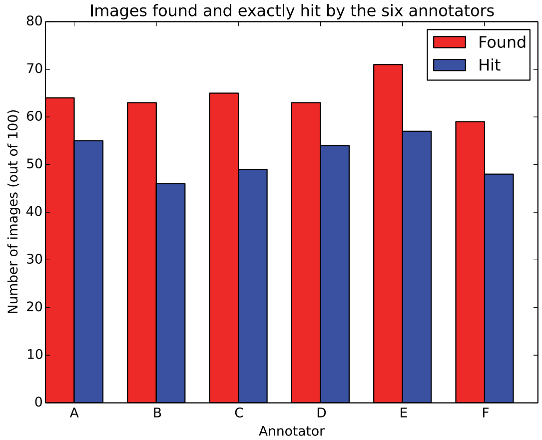
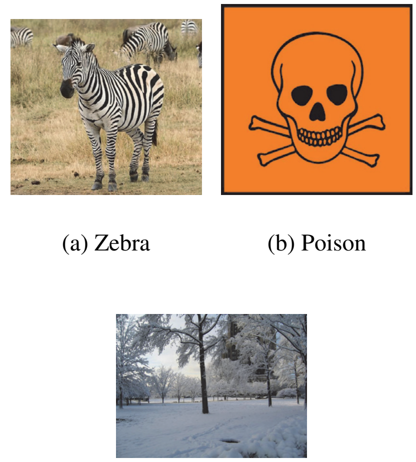


Figure 5: Results of the annotators.



(a) Zebra (b) Poison

(c) Winter

Figure 6: Wikipedia image examples suitable for a word/concept-image correspondence

such as for *poison* (figure 6, (b)), or *sky* also belong the simple correspondences, as well as for example *winter* (figure 6 (c)).

Second, figure 7 shows three words and the corresponding images downloaded from Wikipedia for which no annotator found a correct image. The image for the word *veteran* shows a veteran hospital in Paris. In this case, it is hard to tell that the picture shows a veteran hospital. The image for *nose* shows a tapir. The caption of the image on Wikipedia is "The nose of a tapir". The nose of the tapir is only a small part of the picture and due to the specificity of the nose of the tapir, none of the annotators considered it satisfactory for expressing the concept *nose*. The third picture is perhaps the most obscure: it shows a page from a Soviet calendar, where the image is the first image of the Wikipedia page *week*.

In a third class of cases, there exists a suitable image for a concept symbol on the corresponding Wikipedia page, but it is not ranked as the first one and thus it is not taken in our heuristics. For example, figure 8 shows the three images on the Wikipedia page of *hockey* in ranked order. According to our heuristics, we chose the first image, which was rejected by all six annotators. A simple extension of our image selection method (by including image captions into the ranking process) could help improve the relevance of the results obtained from this class.

Interface integration of the electronic AAC communication board

In our studies for interface integration, we use head-mounted displays to display the communication board. Eye trackers are also used for gaze tracking, symbol selection and initiating utterance/query generation (based on the technical architecture of http://www.dfki.de/RadSpeech/ERmed).

While transferring the communication board to the augmented reality interface and technical infrastructure, we observed that in order to guarantee a flexible and fast interaction with the user, we would highly benefit from personal and large general database collection together with real-time computation in the cloud. The joining of the individual components and the potentials of the suggested approach remain to be explored in the future.

During the progress of our augmented reality developments we have the following wearable interface tools to implement an electronic and ergonomic solution of the adaptable AAC communication board (also cf. figure 1):

- Eye Tracking Glasses (ETG) by SensoMotoric Instruments GmbH: a pair of glasses with eye tracking infrared cameras and a forward looking video camera;
- AirScouter Head Mounted Display (HMD) by Brother Industries: a see-through display which can be attached to the ETG;
- MPU-9150 Motion Processing Unit (MPU) by InvenSense Inc.: a device with integrated motion sensors;
- Epson Moverio BT-100 with MPU: a large viewing angle HMD equipped with MPU;

Several scenarios for AAC have been considered in recent research activities, including gaze-based selection and optical character recognition in real environment (Verő et al. 2014), gaze-based selection of images of the adaptable electronic communication boards projected onto a wall (Vörös et al. 2014), head-motion based selection of images of communicator boards using wide angle displays. Figure 9 shows the hardware scenarios. The Moverio has a broad angle HMD, suitable for presenting a communicator board: gaze restricted OCR can serve context based filtering in real scenarios, e.g., in a convenience store. For a flexible and fast interaction the user needs to search collected personal and large general databases together with real time computation in the cloud. The joining of the individual components and the potentials of the suggested approach remain to be explored and proven.

(a) Nose

(b) Veteran

(c) Week

Figure 7: Some examples where none of the annotators found an appropriate image.

(a) Hockey 1

(b) Hockey 2

(c) Hockey 3

Figure 8: An example where the best suitable image as communication symbol is not the first on the Wikipedia page as here for “Hockey”.

Big Data for Interaction

The first task for flexible and expressive communication is context-based filtering that restricts the number of symbols such that the relevant ones become easily available in a given situation. For example, entering a bistro or a cafeteria food may become the most relevant subject. Furthermore, the backend may know

1. the type of the store and the available selection, e.g., from the GPS coordinates or from the symbols of the store detected by the camera(s) of the smart tool(s);
2. the history of the day, the customs of the user and time that together indicate the goal(s) of the user;
3. the selection of the food types that the user may like, e.g., if the user is vegetarian and/or may not take, e.g., if the user has diabetes;
4. the details of the actual scenario from the camera of the smart tool fused with 9 degree-of-freedom motion sensors and other information sources, including, e.g., optical character recognition (Sonntag and Toyama 2013; Orlosky et al. 2014).

It is relevant to note that our human everyday life is highly regular; ordinary scenarios are easy to forecast. It has been shown (Song et al. 2010) that human actions show a high degree of predictability. This is more so in the case of people with special needs, since their actions and their daily life might be severely limited. However, they are much less protected in unexpected situations and their lives should be guarded by ambient intelligence including information about the environment, route planning and help in decision making. This data can stem wrong automatic and manual access to big data sources. Furthermore, they need fast and easy methods for communication and for the explanation of their intentions; they need online tools for the interaction with other people.

The recognition of ordinary situations and the recognition of larger and unexpected deviations are limited at present, but mathematics is quickly progressing (Candès and Recht 2009; Candès et al. 2011) in this direction. Furthermore, large Internet databases and algorithmic methods that could be used for the fusion of information are becoming available, including (Leich et al. 2013) and some of them are designed to enable machine learning methods to apply, too. Technology is developing quickly, and the interested reader is referred to the Hadoop Summit Series [6]. The execution of the novel mathematical algorithms on large databases and in real time on video streams is on its way, too, see, e.g., (Lohrmann, Warneke, and Kao 2014). As a result, in addition to open-domain (big data) multimedia contents to be accessed by AAC methods in the future, there is a parallel development of real-time parallel multimedia information extraction to allow for real-time interaction and content provision of multimedia material in our adapted communication board.

Conclusions

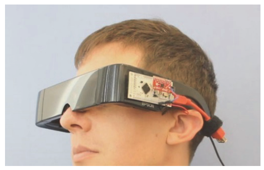
We discussed a dedicated access method for the open-domain QA and IR context and proposed a method for the user to search for additional symbols to be added to the communication board in real-time while using access to big data sources and context based filtering when the desired symbol is missing. We studied the potential extension of the symbol set for representing special events or facts and used a model of human thinking, namely the USF Free Association Norms (Nelson, McEvoy, and Schreiber 1998). We assumed that users can access a symbol similar to the desired one and

[6] http://hadoopsummit.org/

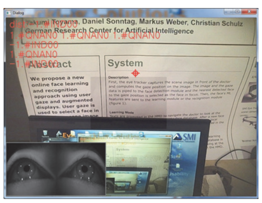
(a) SMI ETG

(b) Brother AirScouter HMD

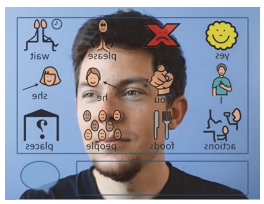
(c) ETG & MD & MPU



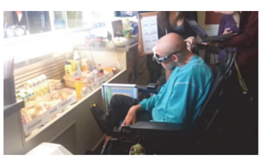
(d) Moverio & MPU



(e) Eye tracking with OCR



(f) Illustration of a 12-symbol board on the Epson Moverio BT-100



(g) Convenience store with gaze tracking and OCR

(h) Gaze based selection on a projected screen

Figure 9: Augmented reality and intelligent user interface tools and scenarios used during the course of our studies.

used Explicit Semantic Analysis (ESA) (Gabrilovich and Markovitch 2007) for finding relevant associated concepts for the communication board. ESA was computed from Wikipedia. We also asked whether the first images of the respective Wikipedia pages provide adequate symbols to be added to the communication board which we evaluated informally. We found that the AAC approach is promising; exceptions concern the history of the concept images or some less representative, possibly culture-dependent reasons for image placement on Wikipedia pages.

We would like to emphasize the need for a better context description for selecting the relevant symbols for the user and for predicting the next ones. This has been the method of choice for, e.g., the popular Dasher tool (Ward and MacKay 2002) that selects the next symbol set according to the probabilities of the letter combinations. In our case, more efficient context-based selection is viable in real life situations where the communication board can be used. Future selection and disambiguation methods of AAC communication boards should should include GPS coordinates, the day and the time, the customs and the needs of the user, including motion patterns and additional sensory information, such as the blood sugar level, for restricting the set of suitable ESA associations by context-dependent board symbols.

Acknowledgements

The research was carried out as part of the EITKIC 12-1-2012-0001 project, which is supported by the Hungarian Government, managed by the National Development Agency, financed by the Research and Technology Innovation Fund and was performed in cooperation with the EIT ICT Labs Budapest Associate partner Group (www.ictlabs.elte.hu).

References

Candès, E. J., and Recht, B. 2009. Exact matrix completion via convex optimization. *Foundations of Computational mathematics* 9(6):717–772.

Candès, E. J.; Li, X.; Ma, Y.; and Wright, J. 2011. Robust principal component analysis? *Journal of the ACM (JACM)* 58(3):11.

Davies, M., and Fleiss, J. L. 1982. Measuring agreement for multinomial data. *Biometrics* 1047–1051.

Demmans, E. C.; Djordjevic, J.; Wu, S.; Moffatt, K.; and Baecker, R. M. 2012. Towards providing just-in-time vocabulary support for assistive and augmentative communication. In *Proceedings of the 2012 ACM International Conference on Intelligent User Interfaces*, IUI '12, 33–36. New York, NY, USA: ACM.

Epp, C. D.; Campigotto, R.; Levy, A.; and Baecker, R. 2011. MarcoPolo: context-sensitive mobile communication support. *Proc. FICCDAT: RESNA/ICTA*.

Gabrilovich, E., and Markovitch, S. 2007. Computing semantic relatedness using Wikipedia-based explicit semantic analysis. In *IJCAI*, volume 7, 1606–1611.

Gabrilovich, E., and Markovitch, S. 2009. Wikipedia-based semantic interpretation for natural language processing. *Journal of Artificial Intelligence Research* 34(2):443.

Kim, G.; Park, J.; Han, M.; Park, S.; and Ha, S. 2009. Context-aware communication support system with pictographic cards. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services*, 86. ACM.

Leich, M.; Adamek, J.; Schubotz, M.; Heise, A.; Rheinländer, A.; and Markl, V. 2013. Applying stratosphere for big data analytics. In *BTW*, 507–510.

Lohrmann, B.; Warneke, D.; and Kao, O. 2014. Nephele streaming: stream processing under qos constraints at scale. *Cluster computing* 17(1):61–78.

Nelson, D. L.; Dyrdal, G. M.; and Goodmon, L. B. 2005. What is preexisting strength? predicting free association probabilities, similarity ratings, and cued recall probabilities. *Psychonomic Bulletin & Review* 12(4):711–719.

Nelson, D. L.; McEvoy, C. L.; and Dennis, S. 2000. What is free association and what does it measure? *Memory & cognition* 28(6):887–899.

Nelson, D. L.; McEvoy, C. L.; and Schreiber, T. A. 1998. The University of South Florida word association, rhyme, and word fragment norms.

Nikolova, S.; Tremaine, M.; and Cook, P. R. 2010. Click on bake to get cookies: Guiding word-finding with semantic associations. In *Proceedings of the 12th International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '10, 155–162. New York, NY, USA: ACM.

Orlosky, J.; Toyama, T.; Sonntag, D.; Sarkany, A.; and Lorincz, A. 2014. On-body multi-input indoor localization for dynamic emergency scenarios: fusion of magnetic tracking and optical character recognition with mixed-reality display. In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2014 IEEE International Conference on*, 320–325. IEEE.

Song, C.; Qu, Z.; Blumm, N.; and Barabási, A.-L. 2010. Limits of predictability in human mobility. *Science* 327(5968):1018–1021.

Sonntag, D., and Toyama, T. 2013. Vision-based location-awareness in augmented reality applications. *LAM Da* 5:2013.

Sonntag, D.; Engel, R.; Herzog, G.; Pfalzgraf, A.; Pfleger, N.; Romanelli, M.; and Reithinger, N. 2007. *SmartWeb Handheld—Multimodal Interaction with Ontological Knowledge Bases and Semantic Web Services*. 272–295.

Sonntag, D.; Reithinger, N.; Herzog, G.; and Becker, T. 2010. *Proceedings of IWSDS 2010—Spoken Dialogue Systems for Ambient Environment*. Springer, LNAI. chapter A Discourse and Dialogue Infrastructure for Industrial Dissemination, 132–143.

Verő, A. L.; Miksztai-Réthey, B.; Vörös, Gy.; Zsigmond, Á.; Pintér, B.; Toyama, T.; Orlosky, J.; Sonntag, D.; and Lőrincz, A. 2014. Gaze tracking and language model for flexible augmentative and alternative communication in practical scenarios. In *16th Biennial Conference of the International Society for Augmentative and Alternative Communication*.

Vörös, Gy.; Verő, A. L.; Pintér, B.; Miksztai-Réthey, B.; Toyama, T.; Lőrincz, A.; and Sonntag, D. 2014. Towards a smart wearable tool to enable people with SSPI to communicate by sentence fragments. In *4th International Symposium on Pervasive Computing Paradigms for Mental Health (MindCare)*.

Ward, D. J., and MacKay, D. J. C. 2002. Fast hands-free writing by gaze direction. *Nature* 418:838.

Wiegand, K., and Patel, R. 2012. Non-syntactic word prediction for aac. In *Proceedings of the Third Workshop on Speech and Language Processing for Assistive Technologies*, 28–36. Association for Computational Linguistics.