Semantically Enriched Cloze Exercice

Technical concept and prototype

Bela-Andreas Bargel, Alexander Streicher, Daniel Szentes, Wolfgang Roller
Department Interoperability and Assistance Systems
Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB
Karlsruhe, Germany
firstname.surname@iosb.fraunhofer.de

Abstract—**In our work we describe an approach for automated and approximate evaluation of character based answers in computer-based tests. Cloze exercises ('fill-in-the-blank') are error-prone to a variety of potential error sources and are subject to syntactic and semantic ambiguities. Our approach combines string similarity measurements with semantic word analysis to render this type of computer-based exercise more robust and user-friendly. We present a working prototype for a web-based e-learning system supporting spelling correction (syntax level) and an extension to synonyms (semantic level). A community-based approach using *OpenThesuarus.org* is proposed to gather and evolve missing synonyms for given answers of cloze exercises.**

e-learning, assessment, cloze exercise, semantic web, open thesaurus, web service

I. INTRODUCTION

While cloze exercises are commonly used in language learning to develop word recognition skills, e.g. [1], they also offer a way to ask for domain specific content created by subject matter experts or technical instructors. The principle of cloze exercises can be used in paper-based assessments as well as in computer-based tests. It has been shown that there is no significant difference between both regarding learning results [2]. Although computer-based tests offer the possibility for automatic evaluation they often lack the capability to distinguish content-related ambiguities and to grasp the correct and indented meaning of the user's input. When a student enters words and/or numbers into a typical text field of an e-learning application, there can be a variety of potential input errors on the syntactic as well as on the semantic level. These errors can be classified by content, characteristics and origin. Our concept describes an integration of different solutions, taking various classes of errors into account.

Imagine the following cloze exercise shown in Fig. 1 inspired by the fictional collie dog character in 'Lassie Come Home' [3]: 'Lassie is a…' followed by an input field. One obvious solution is 'dog' which might be the predetermined solution by the author of this exercise. But also answers like 'dig' (keyboard layout induced mistype for the adjacent 'i' and 'o' keys) or semantically different terms like 'canine', 'collie', 'fictional character' or 'film star' may be considered to be correct in accordance to the defined learning objective of the given exercise. While some answers might be a closer fit to the solution intended by the author ('dog' vs. 'collie') others may seem strange but are still semantically correct in terms of content. In our concept we suggest an interaction scheme where authors of cloze exercises are supported by the system with semantically and content-related alternatives which they can be approve as valid answers. The search space of correct answers is semantically extended by terms which are in synonymous relation to the search query term. In principle this can be regarded as utilizing the concept of a thesaurus to the query.

Figure 1. Cloze exercise example

II. RELATED WORK

In this paper a combination of string similarity measurements with semantic word analysis is suggested to enhance cloze exercises for domain specific use. Cloze exercises have been an active topic of research and method for student assessment for a long time. Taylor showed the cloze procedure in the more general context of comparing reading abilities of different individuals [4]. An application of this has been shown by Guillemette who used it to measure the understandability of an IEEE standard [5]. Cloze exercises have been adopted for various teaching situations as described by Soudek and Soudek [1]. In current computer-based assessment cloze exercises (also referred to as 'fill in the blank') are often realized with syntactic string matching techniques like the Hamming or Levenshtein distance (also called edit distance) [6], [7]. As one of the first general available (freeware) authoring tools the application suite 'Hot Potato' from Half Baked Software offered a user-friendly way to create interactive online exercises like 'multiple-choice, short-answer, jumbled-sentence, crossword, matching/ordering and gap-fill exercises' [8], [9]. Using web based information retrieval for the creation of computer based cloze exercises is shown among others by the project 'Reader-Specific Lexical Practice for Improved Reading Comprehension' [10]. The main idea is to employ a web search engine to find suitable text passages that 'closely reflect the readers' vocabulary and comprehension level to be used to support reading

comprehension and vocabulary growth' [11]. Correia *et al.* recently presented a way for automatic generation of cloze question distractors [12]. In context of semantically enriched cloze exercises this might be enhanced by using semantic relations like synonym, antonym or different meanings as alternative solutions, distractors or assistance for authors while creating cloze exercises. This is also one of the goals of the approach presented in this paper, to enrich cloze exercises by semantic context, to allow alternative meanings and to assist the author.

III. FRAMEWORK

Errors can be classified by content, characteristics and origin (Fig. 2). Each cause can be treated separately and combined in a concept for defect valuation. In our work we consider the following three groups of errors:

- *Typing errors*, for example 'doh' instead of 'dog', as both keys are located side by side. In this case we assume the learners know the correct answer and they are also capable of typing the correct letter but were in some way distracted while typing.
- *Orthographic ambiguities* which lead to different word spelling for the same concept, e.g. in German the word 'dolphin' may be spelled like 'Delfin' (new orthography) or like 'Delphin' (old orthography). Also characters like the German 'ß' incorrectly substituted by 'ss' (but valid in Switzerland) may cause wrong answers in sense of correct spelling but may still be treated as right answers in sense of the intention of the question. Same applies for common spelling differences of British English and American English like in harbour (UK) vs. harbor (US).
- *Incomplete reference*: Valid terms in terms of the given content have not been considered by the author of the exercise and are therefore not taken into account for validation by the e-learning system. Language differences or imprecise questions may lead to semantically correct answers which haven't been considered by the author. Think of 'Lassie is a dog' versus 'Lassie is a collie' (more specific answer) or 'Lassie is a hound' (offering a synonym which on the other hand might also have a negative second association like scoundrel or rascal).

If cloze should be used in domain specific subjects, all these kinds of spelling mistakes and oversights might be treated as didactically irrelevant errors. Considering this assumption synonymously answers may be also treated as correct solutions.

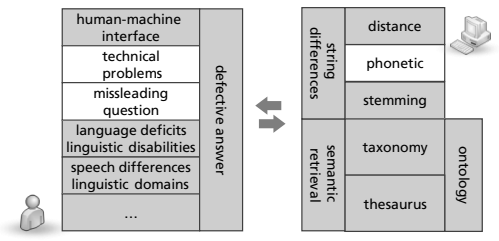

Figure 2. Different causes for errors while answering cloze-exercises

A. Syntax Checking

Syntax errors due to typing errors or different spelling can be handled by computing the distance between the two character strings of the given answer and the reference solution. This feature of proximity is also used in spelling correction algorithms e.g. in word processors or web queries to suggest a replacement for mistyped words.

Common algorithms are the Hamming distance (only for strings of equal length) and the Levenshtein edit distance which may be used for strings with arbitrary size. These techniques are well described [6], [7] and are often implemented in common learning management systems (LMS) like ILIAS [13] or our own e-learning and authoring environment Fraunhofer Crayons [1]. The general idea is to compute the deviation (number of edit operations) of the given and the reference string and to assess whether the solution should be still treated as correct or not according to a given threshold (tolerance level). Common edit operations within the two strings are: '(i) insert a character into a string; (ii) delete a character from a string and (iii) replace a character of a string by another character' [15].

B. Semantic Checking

If a given answer does not pass the check on the syntax level it may still be semantically correct in terms of content. Humans learn to span semantic networks of words and semantic relationships and to search in that network for the true meaning of a word and its context. To check if a given answer is correct as to the contents we can 'search' our networks and decide whether the answer makes sense – or if it is too farfetched to be correct. In information science such a semantic network can be modeled using ontologies [16]. An ontology formalizes the knowledge of a domain and specifies this domain-knowledge by a set of concepts and relationships between those concepts. Hence to build a powerful semantic checking of cloze exercises an ontology can be used to model the domain knowledge of the underlying question. But the problem is that for every question and its domain an ontology engineering process has to take place. This process can be very excessive because domain experts have to specify the words and their meanings and furthermore the relationships between those being specific to the actual context.

To integrate semantically enriched cloze exercises in e-learning systems in a simpler way we focus on a more basic form of an ontology, namely on a thesaurus [17]. A thesaurus may use relationship groups for the linking of terms, namely hierarchical (like in taxonomy), equivalency and associative [18]. To allow synonyms for the semantic cloze the concept of the semantic relation equivalency is set in order. This relation connects similar or near-similar terms. The two hierarchical relations are broader and narrower which specify for a concept if it is more general or rather more specific. This can be seen as the parent/children relationship of a node in a hierarchical tree, where the node represents the term or concept. This semantic relation is called hypernym, or umbrella term, for the parent node and hyponym, or the 'is-a' relation, for the child node.

[1] Crayons® is a platform independent and SCORM compatible e-learning and authoring environment developed by Fraunhofer IOSB [14].

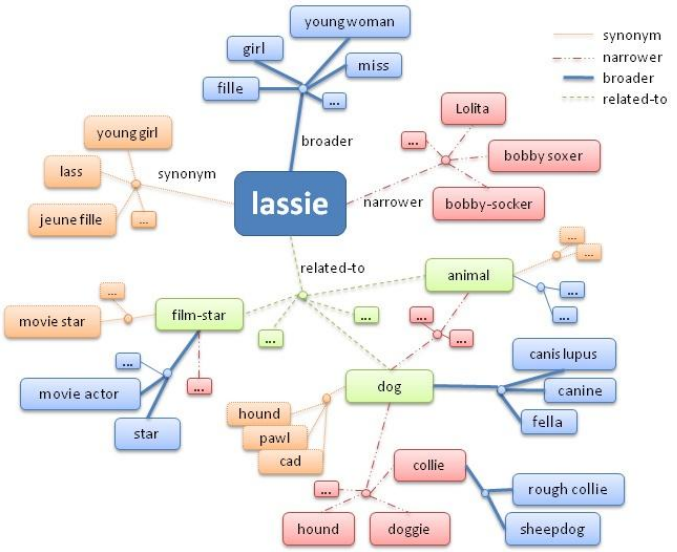

Figure 3. One possible thesaurus network for the term 'lassie'

Fig. 3 shows an example thesaurus network for the term 'Lassie'. The displayed nodes can be seen as valid answers to the cloze exercise 'Lassie is a …'. These semantic relations of a thesaurus extend the principles of simple cloze exercises to semantic cloze exercises where not only the syntax is being checked but also the semantics of the given answers. This can lead the way to intelligent e-learning systems which mimic the expert knowledge of a human tutor. Like human teachers the intelligent e-tutor systems must be able to meet the thinking of the students and accept phrases for the cloze exercises which go beyond a simple text matching and incorporate the meaning of the answer as well.

C. *Didactic use case*

The application scenario of our semantic cloze exercise is the training of Synthetic Aperture Radar (SAR) image interpreters with an e-learning system, in our case the *SAR-Tutor* [19]. The assessment component of such a system can be enhanced by the usage of semantically enriched cloze exercises. The technology of SAR imaging is widely used in various civil as well as defense reconnaissance and surveillance scenarios like the 'Global Monitoring for Environment and Security' (GMES) [20] and includes tasks like digital elevation models, maps, change detection and monitoring of e.g. terrain, pollution, ice layer and vegetation. Radar images offer various advantages over optical images like operational availability under conditions like darkness or cloudiness. However, the task of radar image interpretation is affected by image artifacts, effects and numerous special cases inherent in the system. In context of training different scenarios can be implemented covering and discussing these peculiarities.

For example it might be of interest to discover relevant areas within the image (e.g. moving ground targets) and furthermore to identify their type (e.g. 'compact car', 'mid-range car' or 'trucks'). In this case a hotspot exercise can be combined with a cloze exercise. In the domain of defense there are for example already hierarchical defined categories of objects to support image exploitation and reporting (NATO Standardization Agreements, STANAG) [21]. In favor of motivation and coextensive content these technical terms might be extended by more general synonyms at an early stage of the training. Furthermore relations like broader and narrower are useful when hypernym or hyponym answers are given and the system will be able to assist. This scenario reflects the same requirements as shown in the answer 'Lassie is a collie' instead of 'dog' (*cf.* III.B).

IV. SEMANTICALLY ENRICHED CLOZE EXERCISE – CONCEPT AND PROTOTYPE

A. *Collaborative approach using OpenThesaurus*

To utilize synonyms for cloze exercises a large reference database with both common and domain specific expressions is needed. When designing an e-learning system which should support semantically enriched cloze exercises the effort for the development and maintenance of the underlying knowledgebase must be considered. One way is to focus on synonyms and to exploit existing synonym databases (thesauri) or dictionaries. For the purposes of this work various dictionaries and thesauri have been evaluated for the German language in terms of API access, used data structure, necessary data transfer volume and licensing aspects. This evaluation included OpenThesaurus[2], Wiktionary[3], Woxikon[4], Wortschatz Universität Leipzig[5], Canoo[6] and Duden[7].

We decided to use OpenThesaurus because its community driven approach offers a way to participate and furthermore there is the chance that domain experts from various fields take part in the extension of the thesaurus which eventually endorses the quantitative and qualitative criterion of a meaningful thesaurus source. To do this alone is almost near to impossible and the community driven (social) approach is clearly a reasonable option. OpenThesaurus is an interactive website for the development of a German dictionary of synonyms [23]. It is based on vithesaurus, a web-based thesauri and ontology tool [24], and offers a well-defined API and an XML-based data structure. Vithesaurus *and* the whole dataset of OpenThesaurus are both licensed under the GNU Lesser General Public License (LGPL). Although OpenThesaurus offers no further semantic relations like narrower/broader or is-antonym within its API as of today, the web frontend and the data structure of OpenThesaurus.org already support the valuable metadata association, category, hypernym, hyponym and comment. We expect further use of this metadata when being directly accessible by the API. The dataset of OpenThesaurus is also offered for download as MySQL-dump and text-file. Additionally to general-purpose tools the

[2] OpenThesaurus is a collaborative German open-source thesaurus project, www.openthesaurus.org

[3] Wiktionary is a collaborative lexical database for every language, www.wiktionary.org

[4] Woxikon is a free online dictionary with translations and definitions, www.woxikon.com

[5] Wortschatz Universität Leipzig as vocabulary portal aggregates data automatically from public sources, http://wortschatz.uni-leipzig.de

[6] Canoo offers German dictionaries and grammar, www.canoo.net

[7] Duden is a 'publisher of language reference products […] as well as products for language technology for the German language.' [22].

```
http://www.openthesaurus.de/synonyme/search?q=
        SEARCH-STRING&format=text/xml
```

⇩

```
…
  <matches>
   <metaData> …</metaData>
   <synset id="1234">
     <categories>
       <category name="name of category"/>
      </categories>
     <term term="meaning 1, word 1"/>
     <term term="meaning 1, word 2" level="e.g.
                                colloquially"/>
   </synset>
   <synset id="2345">
     <categories/>
     <term term="meaning 2, word 1"/>
   </synset>
  </matches>
…
```

Figure 1. XML-structure of OpenThesaurus

delimiter-separated values of the OpenThesaurus text-files can be directly explored using the tool OTAccess [25] which also support export of search results as graphs using DOT language (Graphviz) [26], [27].

To query OpenThesaurus.org using its API a specific HTTP GET request has to be sent. The syntax of the request has to be formed like: '*http://www.openthesaurus.de/synonyme/search ?q=SEARCH-STRING&format=text/xml'*. This will result in an XML document which contains data about the search string and the thesaurus synonym data (example in Fig. 4). This structure offers different meanings of the term and corresponding words. The metadata 'level' is used to label register (language level) like technical terminology, colloquially, rough or even vulgar and different German-speaking dialects.

B. Technical integration of OpenThesaurus into a learning management system

There are at least two ways of integrating OpenThesaurus or its data into an e-learning and authoring environment: the first option is to integrate a live online access using the described GET-method inside the authoring environment. The second option would be to download the whole dataset (on a regular basis) and to access the MySQL-dump directly inside the LMS. The second option might be useful as offline option when there is no internet connection for authors available.

As proof of concept the existing cloze exercise in the authoring environment Crayons has been enhanced by the integration of OpenThesaurus. Both options (online link and local dump of the database) have been implemented and its technical realization is the same in most parts. Fig. 7 shows a screenshot of the authoring environment rendered inside a web browser. The buttons labeled 'Synonyms' and 'Synonyms (offline)' will query for corresponding synonyms of the given answer in the 'gap' input field. The drop-down menu with the value '100%' indicates the level of tolerance for the syntax check (*cf.* III.A). In the example of Fig 7 an author has created a cloze exercise and wants to allow further synonyms for the primary reference solution 'freeway'. As the e-learning system is completely web based, the search string 'freeway' is sent from the Crayons front-end client to the web server backend using JSON (Java Script Object Notation) [28]. The implementation of the GUI elements is based upon AJAX technologies and embedding jQuery library functions [29]. In the online version the Crayons backend generates the OpenThesaurus-specific HTTP-GET-string (*cf.* to the example presented above) and sends it to OpenThesaurus.org. The OpenThesaurus server response is an XML-string which is parsed by Crayons to produce a user-friendly presentation output for the GUI. If the local database option is chosen by the author, the workflow is slightly different. Instead of forming the HTTP GET-string, the backend generates the SQL statement for the local copy of the OpenThesaurus database and the response of the local copy is altered directly into the JSON-string without using the XML scheme above.

Fig. 5 picks up the example above: the system offers retrieved synonyms in a tree view. When the author confirms the selected synonyms, the marked words (shown in bold) are sent to the backend and stored in the local database as correct synonyms. In Fig. 6 an example of 'dog' with synonyms and different meanings is shown[8].

In a further step the learning management system should also enable the author to add new synonyms or even new meanings, since most authors will be subject matter experts in their domain. Implementing such a possibility for the local copy of the OpenThesaurus database would be feasible. However the changes would only be local and the community is not able to benefit from the new synonyms. Therefore we suggest an extension of OpenThesaurus to allow uploading of new synonyms using an API.

Figure 4. Synoyms for 'freeway', two selected (in bold)

Figure 5. Synonyms for 'dog'

[8] These examples have been created using an extension of the OpenThesaurus dataset with English words taken partly from Wiktionary.

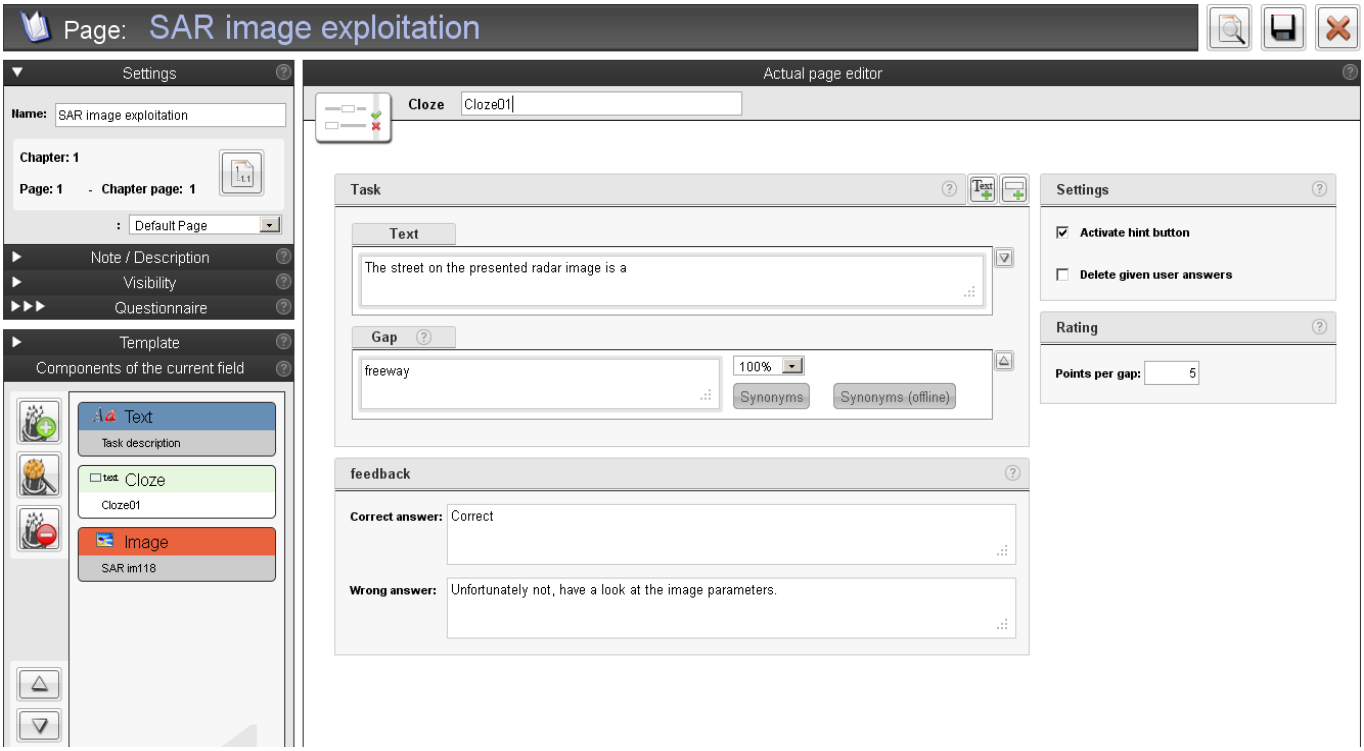

Figure 6. Screenshot of authoring environment (cutout) showing the cloze exercise editor.

V. Conclusion and Outlook

We presented a novel concept and basic principles of a semantically enriched cloze exercise in computer-based tests. The most important aspect thereby is a validation of synonym usage according to subject-didactic and learning objective. If there is a use for synonyms, cloze exercises may be enhanced by alternative solutions.

Collecting these synonyms for the German language may be supported using open thesauri services. One such service is the German thesaurus OpenThesaurus.org which offers a well-defined API and an XML-based document structure. Spelling and typing errors can be recognized and taken into account using phonetic, stemming or distance-based algorithms. Inflections have not yet been regarded but should also be considered in the future. Our prototype successfully implements usage of a synonym web service and the Levenshtein edit distance adjustable for the author of an exercise and offers visual feedback for the learner after evaluation of the given answer: the border of the text input field of a gap is marked in green for fully correct answers and is computed and presented as color gradient into bright red color according to the edit distance and corresponding error level. Dark red is used for incorrect answers.

By implementing a synonym service (online and offline) inside the cloze exercise graphical user interface of an e-learning and authoring environment an evaluation with technical instructors and subject matter experts in the field of synthetic aperture radar image exploitation is planned in future [19]. In this context we expect a number of new synonyms to extent the existing database. According to a community-based approach we suggest adding these to OpenThesaurus either automatically or manually. In this case different communities like OpenThesaurus.org on the one hand and domain specific contributors and users on the other hand would benefit from each other.

The presented concept may also serve as suggestion and strategy for usage of synonyms in other languages inside of cloze exercises focused on domain specific content outside word recognition skills.

Acknowledgment

We would like to thank Daniel Naber of vithesaurus and OpenThesaurus, whom we contacted after our selection of his web service and implementation of our prototype, for his valuable suggestions and for our ongoing discussion about the API of Openthesaurus.org.

References

[1] M. Soudek and L.I. Soudek, “Cloze after thirty years: new uses in language teaching,” *ELT Journal*, vol. 37, Oct. 1983, pp. 335 -340.

[2] M. Li and H. Pu, “Comparison between CBT and PBT: Assessment of Gap-filling and Multiple-choice Cloze in Reading Comprehension,” *Journal of Language Teaching and Research*, vol. 1, Nov. 2010.

[3] F.M. Wilcox, *Lassie Come Home*, movie by Metro-Goldwyn-Mayer Inc. (MGM), Los Angeles, 1943.

[4] W.L. Taylor, “‘Cloze procedure’: a new tool for measuring readability.,” *Journalism Quarterly*, vol. 30, pp. 415-433, 1953.

[5] R.A. Guillemette, “The cloze procedure: assessing the understandability of an IEEE standard,” *IEEE Transactions on Professional Communication*, vol. 32, Mar. 1989, pp. 41-47.

[6] F.J. Damerau, “A technique for computer detection and correction of spelling errors,” *Communications of the ACM*, vol. 7, Mar. 1964, pp. 171–176.
[7] G. Navarro, “A guided tour to approximate string matching,” *ACM Computing Surveys (CSUR)*, vol. 33, Mar. 2001, pp. 31–88.
[8] M. Holmes and S. Arneil, “Hot Potatoes: Free Tools for Creating Interactive Language Exercises for the World Wide Web,” *EuroCALL 1998* (conference presentation), Sep. 1998.
[9] P. Bordeleau, N. Chenik, U. Felix, B. Farrington, H. Hendricks, M. Levy, and K. Walker, “*The History of Computer Assisted Language Learning Web Exhibition*,” Oct. 2000, [Online] Available: http://www.eurocall-languages.org/resources/history_of_call.pdf.
[10] K. Collins-Thompson and J. Callan, “Information retrieval for language tutoring: An overview of the REAP project,” *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*, 2004, pp. 544–545.
[11] K. Dela Rosa, “The REAP Project - Project Description,”*The REAP Project*, Jun. 2011, [Online] Available: http://reap.cs.cmu.edu.
[12] R. Correia, J. Baptista, N. Mamede, I. Trancoso, and M. Eskenazi, “Automatic Generation of Cloze Question Distractors,” *in Proc. Workshop on Second Language Studies: Acquisition, Learning, Education and Technology, Tokyo, Japan*, 2010.
[13] H. Schottmüller, “ILIAS Test & Assessment Documentation,” Jun. 2007, [Online] Available: http://www.ilias.de/docu/data/docu/lm_data/lm_8782/Test_Assessment_38.html.
[14] Fraunhofer IOSB, “*Crayons® - e-learning and authoring environment, product information*,” 2011, [Online] Available: http://www.iosb.fraunhofer.de/?Crayons.
[15] C.D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*, Cambridge University Press, 2008.
[16] T.R. Gruber, “A translation approach to portable ontology specifications,” *Knowledge Acquisition*, vol. 5, Jun. 1993, pp. 199–220.
[17] W. Pidcock, “*What are the differences between a vocabulary, a taxonomy, a thesaurus, an ontology, and a meta-model?*,” Jan. 2003, [Online] Available: http://www.metamodel.com/article.php?story=20030115211223271.
[18] G. Hodge, *Systems of Knowledge Organization for Digital Libraries: Beyond Traditional Authority Files*, Council on Library and Information Resources, 2000, [Online] Available: http://www.clir.org/pubs/reports/pub91/contents.html.
[19] D. Szentes, B.A. Bargel, A. Berger, and W. Roller, “Computer-supported training for the interpretation of radar images,” *EUSAR 2008, 7th European Conference on Synthetic Aperture Radar 2008*, Berlin: VDE-Verlag, 2008.
[20] European Commission, “GMES - Observing our planet for a safer world - Space - Enterprise and Industry,” Jul. 2011, [Online] Available: http://ec.europa.eu/enterprise/policies/space/gmes/.
[21] NATO, “*AEDP-2 - NATO Intelligence, Surveillance, and Reconnaissance Interoperability Architecture(NIIA)*,” 2005, [Online] Available: http://www.nato.int/docu/stanag/AEDP02/aedp2_home.htm.
[22] M. Alexa, B. Kressig, M. Liepert, K. Reichenberger, L. Rostek, K. Rautmann, W. Scholze-Stubenrecht, and S. Stoye, “The Duden Ontology: An Integrated Representation of Lexical and Ontological Information,” *LREC Workshop on WordNet Structures and Standardisation, and How These Affect Wordnet Applications and Evaluation*, Las Palmas, Gran Canaria, 2002.
[23] D. Naber, “*OpenThesaurus: Building a Thesaurus with a Web Community*,” Mar. 2004, [Online] Available: http://www.openthesaurus.de/download/openthesaurus.pdf.
[24] D. Naber, “*vithesaurus | Download vithesaurus software for free at SourceForge.net,”* Mar. 2009; [Online] Available: http://sourceforge.net/projects/vithesaurus/.
[25] T. Zastrow, “*OTAccess - Java-GUI für OpenThesaurus*,” Apr. 2009, [Online] Available: http://www.thomas-zastrow.de/projekte/ot/manual.pdf
[26] E. Gansner and S. North, “An open graph visualization system and its applications to software engineering,” *Software - Practise and Experience*, vol. 30, 2000, p. 1203--1233.
[27] E. Gansner, E. Koutsofios, and S. North, “*Drawing graphs with dot - dot User’s Manual*,” AT&T Bell Laboratories, Murray Hill, NJ, 2006.
[28] Douglas Crockford, “*The application/json Media Type for JavaScript Object Notation (JSON)*,” Jul. 2006, [Online] Available: http://tools.ietf.org/html/rfc4627.
[29] J. Chaffer and K. Swedberg, *jQuery Reference Guide: A Comprehensive Exploration of the Popular JavaScript Library*, Packt Publishing, 2007.