The Dicta-Sign Wiki: Enabling Web Communication for the Deaf

Eleni Efthimiou[1], Stavroula-Evita Fotinea[1], Thomas Hanke[2], John Glauert[3], Richard Bowden[4], Annelies Braffort[5], Christophe Collet[6], Petros Maragos[7], and François Lefebvre-Albaret[8]

[1] Institute for Language and Speech Processing/ATHENA RC, Ahens, Greece
{eleni_e,evita}@ilsp.gr
[2] Universität Hamburg, Hambourg Germany
thomas.hanke@sign-lang.uni-hamburg.de
[3] University of East Anglia, Norwich, United Kingdom
J.Glauert@uea.ac.uk
[4] University of Surrey, Surrey, United Kingdom
R.Bowden@surrey.ac.uk
[5] LIMSI/CNRS, Orsey, France
annelies.braffort@limsi.fr
[6] Université Paul Sabatier, Toulouse, France
collet@irit.fr
[7] National Technical University of Athens, Ahens, Greece
maragos@cs.ntua.gr
[8] WebSourd, Toulouse, France
francois.lefebvre-albaret@websourd.org

Abstract. The paper provides a report on the user-centred showcase prototypes of the DICTA-SIGN project (http://www.dictasign.eu/), an FP7-ICT project which ended in January 2012. DICTA-SIGN researched ways to enable communication between Deaf individuals through the development of human-computer interfaces (HCI) for Deaf users, by means of Sign Language. Emphasis is placed on the Sign-Wiki prototype that demonstrates the potential of sign languages to participate in contemporary Web 2.0 applications where user contributions are editable by an entire community and sign language users can benefit from collaborative editing facilities.

Keywords: Sign language technologies, sign-Wiki, multilingual sign language resources, Deaf communication, Deaf user-centred HCI.

1 Research Rational

The idea driving research in DICTA-SIGN was that development of Web 2.0 technologies have made the WWW a place where people constantly interact with each other by posting information (e.g. blogs, discussion forums), modifying and enhancing other people's contributions (e.g. Wikipedia), and sharing information (e.g. Facebook, social

K. Miesenberger et al. (Eds.): ICCHP 2012, Part II, LNCS 7383, pp. 205–212, 2012.

news sites). However, these technologies are not friendly to sign language users, because they require the use of written language. Sign language (SL) videos cannot fulfil the same role as written text in these new technologies, given, first, that they are not anonymous and, second, that people cannot easily edit and add to a video which someone else has produced. Under this light, DICTA-SIGN's goal was to develop the necessary technologies that make Web 2.0 interactions in sign language possible. To fulfil its goals, DICTA-SIGN dealt with four sign languages: British Sign Language (BSL), German Sign Language (DGS), Greek Sign Language (GSL) and French Sign Language (LSF). The project involved research from several scientific domains in order to develop technologies for sign recognition and generation, exploiting significant knowledge of the structure, grammar and lexicon of the project Sign Languages.

2 Technological Advancements and Systems Integration

DICTA-SIGN contributed new knowledge in the domain of SL resources acquisition and corpus construction but also in the front of SL technologies, advancing research in image processing, computer vision, the statistical methods for continuous sign recognition with multimodal fusion and adaptation, as well as in virtual human technology. Another significant contribution, beyond advances in related technologies, is the user-centred interface design of the project's prototypes, which derived from systematic involvement of end-users in evaluation procedures accompanying development phases. All technologies and resources developed during the project were integrated in a sign language wiki, which enabled not only to showcase the project's outcomes, but also allowed for extensive evaluation and actual use by end-users. The prototype is based on the initially set scenario, where the user signs to a webcam providing as input isolated signs or continuous signing. The computer recognizes the signed phrases, converts them into an internal representation of sign language, and then has an animated avatar sign them back to the user. Content on the Web is then contributed and disseminated via the signing avatar.

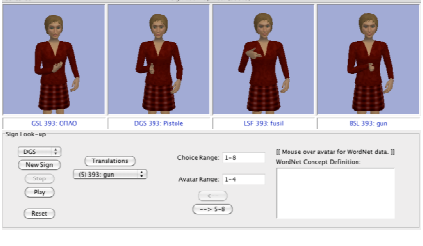

Fig. 1. Sign Look-up Tool results

The required linguistic knowledge was acquired on the basis of creation and annotation of a specially designed parallel corpus [1] and a related multilingual lexical database. The project language resources were primarily used in order to extract generalised lexical and grammar SL models, to support sign language recognition [2-6]

and animation technologies [7]. The same resources were also exploited in order to allow for the application of a simple sign-to-sign translation service, integrated into the project's Sign-Wiki.

As expected, a major effort has been placed on systems' integration linking recognition technology to synthesis and animation, both supported by adequately annotated SL resources [8]. An initial Search-by-Example interface [9] to the DICTA-SIGN lexical database has been enhanced to produce a Sign Look-up Tool (Fig. 1) that provides users with a simple sign-level translation tool for exploring corresponding signs in the four project sign languages.

The Sign Look-up Tool enables a Deaf user to perform a sign and see the corresponding signs in the four project languages. The system uses an Xbox Kinect™ device to recognise signs. The sign look-up tool plays back the recognised sign, or the closest matches, to the user using an avatar. The user can then see the corresponding sign in all four languages.

Fig. 2. Tracking of (a) corpus data and (b) facial features

A practical use for this tool is where a sign language user sees a sign in an unfamiliar language on a video or when travelling to another country. Successful implementation of the Sign Look-up Tool led to integration plans for the sign language Wiki. With significant enhancement of the sign recognition technologies (Fig. 2a, 2b, 3) the sign language Wiki prototype was identified as the most useful demonstration of DICTA-SIGN technologies for Deaf people.

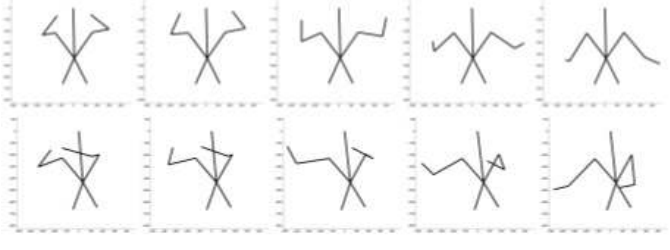

Fig. 3. Demonstration of co-articulation-relocation of signs for the continuous sign recognition framework: In the example, each row corresponds to an articulation instance of the sign TABLE which are articulated in continuous mode and located in different positions. Results are shown via the Skeleton data captured with the Kinect.

3 A Sign Language Wiki

A major requirement of contemporary Web 2.0 applications is that user contributions are editable by an entire community. The oldest, and most popular, application of this type is a Wiki, where any contribution can be edited and refined, anonymously if so wished, by someone else. As the success of Wikipedia and related sites show, this type of community collaboration results in a rapid amassing of knowledge.

There is no doubt that sign language users could benefit similarly from collaborative editing. With this in mind, a server was developed to provide the same service as a traditional Wiki, but using sign language. Instead of using text as the output medium, a signing avatar presents information (fig. 4). The use of an avatar preserves the anonymity of the user, and facilitates modification and reuse of information present on the site.

A typical Sign-Wiki page is divided in two sections, a Content Structure panel on the left to accommodate content and a Content Presentation panel on the right, where the signing avatar appears.

The Content Structure panel presents a structured view of the page content to the user allowing for navigation within the page and selection of units for presentation by the virtual human signer.

The global level of detail at which the content page structure is displayed can be controlled by the user by appropriate selection/deselection of the nodes in a tree structure which may correspond to individual signs or sign phrases, while each individual sign is labelled with its spoken language gloss name.

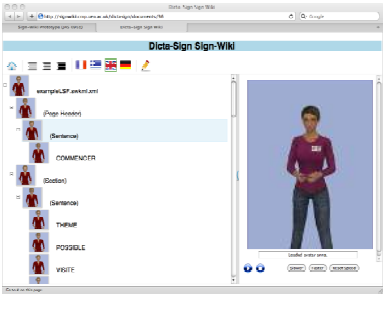

Fig. 4. Standard Sign-Wiki page layout

To view the page content, or any structural unit within the page using the virtual human signer, the user simply selects by clicking on the required node, so as the virtual human signer initiates presentation of the selected sequence sign by sign. As each individual sign is presented, its entry in the tree structure on the left is highlighted in blue, thus providing visual feedback to the user, also allowing for verification of the structural organisation of the content.

For any given Sign-Wiki content page, a switch from Viewing mode to Edit/Input mode is achieved by simple clicking of the relevant icon. In Edit/Input mode the user may execute the following actions:

- Save updates to the page at any stage during editing.
- Search the lexical resources of the DICTA-SIGN lexicon for a sign or signs to be added to the content of the page under editing (Fig. 7).
- Interact with the system to directly input signs, via a Kinect tracking device.

In Edit/Input mode the Content Structure panel gains the ability to act as a Drag-and-Drop based structure editor. The user can select any node within the content tree structure and move it to a new position in the tree by dragging it there, while it is possible to view any unit within the structure as often as the user desires while editing proceeds.

The currently available sign or sign sequence insertion methods involve Kinect input by exploitation of Sign Recognition technologies, Search in the Dicta-Sign Lexicon to enrich the nodes in the Content Structure panel, and an individual sign building functionality. Both the first and second of these methods allow identification of signs, if already included in the Dicta-Sign lexicon. The sign building environment allows the user to create a sign definition by identifying its individual features (hand-shape, orientation, location, movement, etc.) interactively. It also allows the user to edit signs already included in the page, using the same set of interactive techniques.

Fig. 5. Sign-Wiki input reviewing environment

The system acts as a dictation machine using sign, providing recording, playback (Fig. 5), and editing (Fig. 6). As regards Sign-Wiki implementation issues, the prototype is built using conventional Web technology, augmented with various special purpose SL processing modules. On the server side, the Sign-Wiki Web site is implemented using the Django framework, supported by a PostgreSQL database. The client-side of the Web site is implemented using HTML and Javascript, supported by the JQuery library, and several other small JQuery-related library modules. The system also integrates a dedicated Kinect Input/Recognition server.

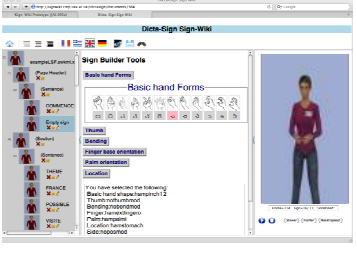

Fig. 6. Sign-Wiki sign building environment

To use the Sign-Wiki the user needs just a conventional personal computer system and a web browser (Windows or Mac OS X, with Firefox). To do Kinect-based SL input, a Microsoft Kinect™ input device is required, attached either to the client computer itself or to another local computer, physically located close the main client system, and connected via the local network to that system. The local computer hosting the Kinect device also needs to have the special Kinect-service software, developed within the project, installed on it. Apart from this, the Sign-Wiki depends only on standard client-side Web software technology.

Innovation achieved by DICTA-SIGN in respect to communication via sign language is noticeable in relation to the pre-DICTA-SIGN state-of-the-art regarding a number of barriers causing parameters which mainly relate to the exclusive use of video to convey the signed linguistic message. For the first time technological solutions provide the conditions so that a signing user may create new SL content by either entering his/her own productions to the system by means of a Kinect device and/or using a set of sign creation tools which are also used for editing purposes. The Web user may edit previously uploaded signs or sign phrases by applying i.e. simple copy-paste procedures on pieces of SL utterances or by changing basic components of a sign, using a visual sign editor or the set of HamNoSys notations [10]. Sign-Wiki users, like any other Wiki users, may view SL information uploaded by other individuals. This may involve information in one's own sign language or may require translation support in order to be comprehended. In the latter case, the user may find support by a Sign-Look-up translation module, which currently allows search of signs in four sign languages. This facility is rather supportive in the multilingual Web environment, since multilingual correspondences of the same concept increase the possibility of its understanding. Finally, answering a demand been strongly expressed by the Deaf communities, the user may save, upload and present his/her content preserving his/her anonymity, since performance of sign language content takes place by means of a signing avatar.

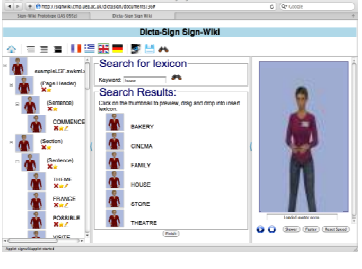

Fig. 6. Sign-Wiki lexicon search results

The DICTA-SIGN prototypes have been exposed to end-user evaluation procedures that have provided comments relating to all levels of implementation, crucially emphasising on the Deaf users' preferences in respect to interaction with the systems, thus, gaining advanced human-computer interface design for Web 2.0 sign language applications, which can be best viewed in implementation of the Sign-Wiki prototype. Especially in respect to the Sign-Wiki, since the prototype is usable online, all

functions were tested via internet by end-users using one of the four project sign-languages (LSF, GSL, DGS, GSL) thanks to the translation option. Gained results revealed that the Wiki is actually used equally in order to create new utterances and to modify existing utterances. While it would also be possible to use the Wiki interface key concepts in pedagogical applications or for information providing purpose in combination with other existing solutions like 3DSigner (www.3DSigner.fr), besides possible applications, the testers pointed out provided anonymity as the major strength of such an application. Detailed reporting on end-user evaluation of the DICTA-SIGN Sign-Wiki, is the subject of project deliverable *D8.2: Evaluation report of Sign-Wiki demonstrator*[1].

4 Conclusion

DICTA-SIGN has undertaken fundamental research and development in the combined use of image processing and advanced computer vision techniques, statistical methods for continuous sign recognition with multimodal fusion and adaptation, virtual human technology, sign language modelling, grammar & lexicon design and development as well as corpus construction. The DICTA-SIGN demonstrator focused on the end user's requirements as regards human-computer interaction via sign language. Under this light, the main aim here has been to underline the range of actions and interaction possibilities that are finally offered to signing users of Web 2.0, resulting from research work that exploits properly annotated language resources.

Acknowledgments. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 231135.

References

1. Matthes, S., Hanke, T., Storz, J., Efthimiou, E., Dimiou, N., Karioris, P., Braffort, A., Choisier, A., Pelhate, J., Safar, E.: Elicitation Tasks and Materials designed for Dicta-Sign's Multi-lingual Corpus. In: Dreuw, P., Efthimiou, E., Hanke, T., Johnston, T., Martínez Ruiz, G., Schembri, A. (eds.) LREC 2010, Proc/ 4th Workshop on Representation and Processing of Sign Languages: Corpora and Sign Language Technologies, pp. 158–163 (2010)
2. Ong, E.-J., Bowden, R.: Robust Facial Feature Tracking using Shape-Constrained Multi-Resolution Selected Linear Predictors. IEEE Trans. Pattern Analysis and Machine Learning 33(9), 844–1859 (2011)
3. Cooper, H., Pugeault, N., Ong, E.-J., Bowden, R.: Sign Language Recognition using Sub-Units. J. of Machine Learning and Recognition: Gesture Recognition (to appear, 2012)

[1] Available at: `http://www.dictasign.eu/attach/Main/PubliclyAvailable ProjectDeliverables/DICTA-SIGN_Deliverable_D8.2.pdf`

4. Pitsikalis, V., Theodorakis, S., Vogler, C., Maragos, P.: Advances in Phonetics-based Sub-Unit Modeling for Transcription Alignment and Sign Language Recognition. In: Proc. IEEE CVPR Workshop on Gesture Recognition, Colorado Springs, USA (2011)
5. Theodorakis, S., Pitsikalis, V., Maragos, P.: Advances in Dynamic-Static Integration of Manual Cues for Sign Language Recognition. In: Proc. 9th International Gesture Workshop (GW 2011): Gesture in Embodied Communication and Human-Computer Interaction, Athens, Greece (May 2011)
6. Theodorakis, S., Pitsikalis, V., Rodomagoulakis, I., Maragos, P.: Recognition with Raw Canonical Phonetic Movement and Handshape Subunits on Videos of Continuous Sign Language. In: Proc. Int'l Conf. on Image Processing (ICIP 2012), Orlando, Florida, USA (September 2012) (submitted)
7. Jennings, V., Elliott, R., Kennaway, R., Glauert, J.: Requirements for a Signing Avatar. In: Proc. Workshop on Corpora and Sign Language Technologies (CSLT), LREC 2010, Malta, pp. 33–136 (2010)
8. Efthimiou, E., Fotinea, S.-E., Hanke, T., Glauert, J., Bowden, R., Braffort, A., Collet, C., Maragos, P., Lefebvre-Albaret, F.: Sign Language technologies and resources of the Dicta-Sign project. In: Proc. of the 5th Workshop on the Representation and Processing of Sign Languages: Interactions between Corpus and Lexicon. LREC 2012, May 23-27, Istanbul, Turkey (to appear, 2012)
9. Elliott, R., Cooper, H., Ong, E., Glauert, J., Bowden, R., Lefebvre-Albaret, F.: Search-By-Example in Multilingual Sign Language Databases. In: Proc. Sign Language Translation and Avatar Technologies Workshops (2011)
10. Hanke, T.: HamNoSys - representing sign language data in language resources and language processing contexts. In: LREC. Workshop Proceedings, Representation and Processing of Sign Languages, pp. 1–6. ELRA, Paris (2004)