Sign Language technologies and resources of the Dicta-Sign project

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, [2]Universität Hamburg, [3]University of East Anglia, [4]University of Surrey, [5]LIMSI/CNRS, [6]Université Paul Sabatier, [7]National Technical University of Athens, [8]WebSourd

[1]{eleni_e,evita}@ilsp.gr, [2]thomas.hanke@sign-lang.uni-hamburg.de, [3]J.Glauert@uea.ac.uk, [4]R.Bowden@surrey.ac.uk, [5]annelies.braffort@limsi.fr, [6]collet@irit.fr, [7]maragos@cs.ntua.gr, [8]francois.lefebvre-albaret@websourd.org

Abstract

Here we present the outcomes of Dicta-Sign FP7-ICT project. 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. It has researched and developed recognition and synthesis engines for sign languages (SLs) that have brought sign recognition and generation technologies significantly closer to authentic signing. In this context, Dicta-Sign has developed several technologies demonstrated via a sign language aware Web 2.0, combining work from the fields of sign language recognition, sign language animation via avatars and sign language resources and language models development, with the goal of allowing Deaf users to make, edit, and review avatar-based sign language contributions online, similar to the way people nowadays make text-based contributions on the Web.

Keywords: Sign language technologies, multilingual sign language resources, sign language Web 2.0 applications, Deaf communication, human-computer interfaces (HCI) for Deaf users

1. The Dicta-Sign framework

When introducing the components and objectives of the Dicta-Sign project (FP7-ICT) at the 4th Workshop on the Representation and Processing of Sign Languages in the framework of LREC-2010 (Efthimiou et al., 2010), we underlined its major aim, which was to improve the state of web-based communication for Deaf people by allowing use of sign language in a way similar to what holds for oral languages. In the three years of its life-cycle, Dicta-Sign significantly progressed knowledge in the domains of recognising and generating authentic signing. It also created significant sign language resources to be exploited in future research both in sign language linguistics and sign language technologies. Dicta-Sign developed several technologies demonstrated via a sign language-aware Web 2.0 application, a wiki, combining work from the fields of sign language recognition, sign language animation via avatars and sign language linguistics. The goal is to allow Deaf users to make, edit, and review online sign language content presented via a signing avatar, in a similar way to how people make text-based contributions on the Web. In this context, several successful proof-of-concept demonstrators have been created. Furthermore, to serve its technological goals, Dicta-Sign created a multilingual lexicon and annotated comparable corpora resources for four European sign languages: Greek, British, German, and French Sign Language, which were exploited in the definition of models for sign language lexicon and grammar, they informed the project's technologies and are now available to the international research community.

Dicta-Sign (http://www.dictasign.eu/) actually developed three proof-of-concept prototypes: a Search-by-Example Tool that works as a sign language dictionary (Cooper et al,. 2011), a Sign Look-up Tool that allows for multilingual translation of signs in the four project sign languages (Elliot et al., 2011), and a Sign-Wiki, a sign language-based Wiki which enables end-users to view, edit, create and upload sign language content, while preserving their anonymity through the use of a signing avatar for the display of so created content.

2. Technological advancements and multilingual SL resources

Dicta-Sign contributed new knowledge in regard to sign language resources creation and the whole range of sign language technologies. The scientific domains addressed are:

- Image processing
- Advanced computer vision
- Statistical methods for continuous sign recognition with multimodal fusion and adaptation
- Virtual human technology
- Sign language modelling
- Grammar and lexicon design and development
- Corpus construction

The project has investigated sign language processing in different input and output modalities Within these input and output modalities, different channels (manual signing, eye gaze, facial expression and body posture) present complementary communicative information. Dicta-Sign has analysed and synthesised information within these channels to encompass the full range of expressiveness within sign languages.

In so doing, Dicta-Sign has achieved the following goals:

- Establishment of the world's only extensive parallel multilingual corpus of annotated sign language video data to inform further sign language research and provide a resource to inform sign language processing technology (http://www.sign-lang.uni-hamburg.de/dicta-sign/portal/).
- Development of advanced sign language annotation tools to provide access to this corpus integrating the

recognition, generation and animation technologies developed within the project (Dicta-Sign technical report: D5.4; Gonzalez & Collet, 2012)

- Use of image processing and computer vision technology and the development of a statistical framework to enhance and progress beyond the state-of-the-art in continuous sign language recognition, also experimenting with sign language linguistic models and exploiting multimodal fusion and adaptation (Dicta-Sign technical reports: D1.3. and D2.3)
- Development and dissemination of a morphophonetically based orthography based upon HamNoSys incorporating the manual and non-manual channels of sign language communication used both for annotation purposes and as an intermediate representation for both sign language generation and recognition within system software,
- Provision of large cross-lingual sign lexicons (http://www.sign-lang.uni-hamburg.de/dicta-sign/consign/demo/cs_list_eng.html),
- Extension of sign language generation and virtual human animation technology to advanced realistic signing (Dicta-Sign technical report: D3.3),
- Development of a prototype sign-to-sign translation service (Dicta-Sign technical report: D7.2).

These objectives are structured through the development and delivery of the following outputs:

- A parallel multilingual corpus for four national sign languages – German, British, French and Greek (DGS, BSL, LSF and GSL respectively).
- A multilingual lexicon of 1000+ signs from each of the four project sign languages,
- A continuous sign language recognition system that has achieved significant improvements and also has researched the novel directions of multimodal sign fusion and signer adaptation,
- A language generation and synthesis component, covering in detail the role of manual, non-manual and placement within signing space,
- Annotation tools which incorporate these technologies providing access to the corpus and whose long term utility can be judged by the up-take by other sign language researchers,
- Three bidirectional integrated prototype systems which showcase the utility of the system components beyond the annotation tools application.

All project prototypes have been exhaustively evaluated by Deaf end users. Their details, emphasising on their relation to project resources, are presented in section 3 below.

2.1 Sign language recognition

In order to support work in sign recognition, from its early stages, Dicta-Sign has created 3D data of sign language by shooting the whole corpus with Bumblebee depth capturing cameras, which were incorporated in the studio set up used for the project corpus acquisition. With the arrival of the Kinect, this new device was adopted as a 3D data acquisition instrument and Kinect data have since fed experimentation in both visual tracking and feature extraction as well as in continuous sign recognition research.

2.1.1. Visual tracking and feature extraction

In order to recognise isolated signs, a low level feature description of the signers' actions need to be extracted. Visual tracking and feature extraction work in Dicta-Sign has covered several tasks including hand detection and tracking in 2D and 3D (Hadfield & Bowden 2011 and 2012), visual hand detection emphasising on tracking and motion feature extraction (Roussos et al., 2010a, Roussos et al., 2010b), body part detection with estimation from depth (Holt et al., 2011), facial feature localisation and tracking (Ong & Bowden, 2011), face modelling for tracking and related feature extraction (Rodomagoulakis et al. 2011; Antonakos et al. 2012), handshape modelling and related feature extraction (Rousos et al. 2010a; Cooper et al., 2012, Pugeault 2011), as well as mapping of body motion to HamNoSys features.

2.1.2. Continuous sign language recognition

For recognition of dictation style sign the project developed a new spatio-temporal feature selection approach called Sequential Pattern Trees (Ong et al,. 2012) which generalises well to unseen signers and forms the heart of the Sign Wiki recognition module. For continuous sign language recognition, the project developed an efficient visual front-end for spatio-temporal processing of the corpus video data and statistical (HMM-based) models for both data-driven and phonetics-based Sub-units (SUs) which have allowed Dicta-Sign to make significant advances in continuous SL modelling and recognition. Specifically, the project's contribution to new knowledge in continuous sign recognition includes: 1) dynamic vs. static data-driven SUs (Pitsikalis et al. 2010, Theodorakis et al. 2011a), 2) phonetics-based SUs (Pitsikalis et al. 2011; Vogler 2011), 3) raw and canonical SU models (Theodorakis et al. 2012, Maragos et al. 2012), which provide novel solutions in case of multiple signers and system adaptation to them, 4) facial event detection and recognition of related linguistic cues (Antonakos et al. 2012), and 5) fusion of multiple cues/modalities for improving continuous sign recognition (Roussos et al. 2010, Theodorakis et al. 2011a, Theodorakis et al. 2011b, Theodorakis et al. 2012). In addition, exploitation of SL grammar rules has also led to interesting results on continuous video data.

2.2 Sign language generation and animation

Dicta-Sign has improved the state-of-the-art in synthesis and animation of sign language through 3D virtual characters (Glauert & Elliott, 2011). Synthesis is based on the use of SiGML, Signing Gesture Markup Language, which is an XML language based heavily on HamNoSys, the Hamburg Notation System, since HamNoSys

transcriptions can be mapped directly to SiGML. SiGML is primarily implemented through the JASigning software which is supported on both Windows and Macintosh platforms. During the project, enhancements have been made to SiGML that allow precise control of the timing of animations (Jennings et al., 2010). This was exploited to show that an animation synchronised with a sign language video can be produced by annotating the signs used, along with their timings, and playing back HamNoSys transcriptions of the signs (Hanke et al., 2011), where fine control of the location of sign postures and the details of transitions between postures has been provided.

2.3 SL linguistic modelling

Work on linguistic modelling was concerned with the high-level linguistic framework relevant to sign languages. The aim was to develop lexicon and grammar models to assist in SL generation and recognition.

On lexical level, the proposed models provide formal descriptions that may be used by a sign recognition system, whose input is a video of signed utterances, or as input to an animation system, whose output is signed by an animated virtual signer. To do so, these models need to include both phonetic representations based on HamNoSys and on the Johnson & Liddell phonetic notation, as well as abstract level representations based on Zebedee[1], a sign description model based on a geometric tool kit and time structuring that allows signs and their variability in the signing space to be specified.

Regarding grammar modelling, the main issues include flexible sign order in sign phrase, signing space representation, non-manual gestures, synchronisation of multiple body articulators, role shift, and prosody (Filhol 2012). Grammar modelling work in Dicta-Sign concentrated on a set of phenomena which provide cues to all above:

1. Enumeration: sets of unordered elements signed in sequence;
2. Alternatives, where options of a choice are listed sequentially;
3. Qualification/naming, where one sign in a sequence denotes an entity specified (named, finger-spelt, adj-qualified) by the others (Filhol & Braffort 2012);
4. Neutral questions, i.e. the case where the speaker is directly asking for an answer which he is not able to predict;
5. Quantifiers (small/big);
6. Announcing titles: announcing the topic of a discourse section.

The first three in the above list cover utterance components structure, the fourth one deals with the whole clause structure, the fifth affects lexical units, while the last one touches upon discourse structure.

2.4 Sign language annotation tools

Concerning annotation tools, Dicta-Sign proposed a complete solution that answers the need to integrate various tools into the annotation process.

The proposed solution is specified as a distributed system architecture called the Automatic Annotator Assistants System (Collet, Gonzalez, Milachon 2010). It aims to use software developed in different programming languages, operating systems and software platforms, outsourcing annotation tasks to other machines such as a computing cluster, easily adding automatic processing to support the annotation task within existing annotation tools.

This architecture is multi-platform and multi-language (including RealBasic, C/C++ and Java) and the model used for data exchange is adaptable to a number of annotation formats. It also includes a security mechanism to ensure annotation data, video files and program code integrity (Dubot & Collet 2012). As a prove of concept, we have implemented the following automatic annotation assistants (A3): Body Part Tracking (Gonzalez & Collet, 2011); Facial Feature Tracking (Ong & Bowden, 2011); Signing Detection; Sign Segmentation (Gonzalez & Collet, 2012). All specifications of API for A3 and the source code of A3 template and A3S are publicly available.

2.5 The Dicta-Sign sign language resources

Within the Dicta-Sign project sign language resources were compiled for four European languages: British, French, German, and Greek Sign Language. These resources were used to inform progress in other research areas within the project, especially sign recognition, linguistic modelling, and sign generation.

A multilingual lexical database providing a core lexicon of approximately 1000 entries in the four project sign languages was the first resource to be built. The shared list of concepts chosen for the lexicon is of everyday use or specifically related to the field of Dicta-Sign's main topic, European travel. Signs were recorded for each language and annotated by assigning gloss labels, form description (HamNoSys) and a rough meaning. However, the biggest achievement in the area of language resources within the project is Dicta-Sign's multilingual corpus on the domain "Travel across Europe". Prior to the project, parallel corpus collection for sign languages had only been undertaken in minimal sizes or for spoken language simultaneously interpreted into several sign languages, but not for semi-spontaneous signing by native signers, based on a well designed methodology for corpus elicitation, which eliminates dominant oral language interference (Matthes et al. 2010).

Data collection took place in all four countries involved in the project, using seven different cameras to film the informants from different perspectives (front, side, and top view) as well as with additional stereo cameras that provide video footage for automatic processing. In each country, 14 to 16 informants were filmed in sessions lasting about two hours each. This resulted in applicable language data of at least 8 hours per language.

After various steps of video post-processing the data were annotated using the iLex annotation environment (Hanke

[1] http://perso.limsi.fr/filhol/zebedee/index.html

& Storz 2008).
The detailed annotation, conducted for parts of the corpus data, includes segmentation of the continuous signing into individual signs, lemmatisation (i.e. assigning glosses), a form description of the signs using HamNoSys as well as English translation. Additionally, content tags were assigned to most of the corpus data that reflect the topics the informants signed about. These tags allow finding video sections with comparable content across individual informants and languages.
The Dicta-Sign language resources are made available via a dedicated web portal that provides data on different access levels and with different approaches to access the data. The access levels range from publicly available data to restricted access for researchers or future project partners.
Different approaches are offered to access the data, e.g. by selecting a certain language and informant, by choosing a certain task, or by selecting a specific content tag to see data of various languages and informants signing about the same topic.

3. Showcase prototypes

The main objectives of Dicta-Sign concentrated towards the development of an integrated framework that allows contributions in the project's sign languages. The realisation of this goal is crucially based on the assumption that Deaf users may directly insert their own content via Kinect devices. This content is recognised by the sign language recognition component and converted into a linguistically informed internal representation, which is used to animate the user's contribution with an avatar, and/or to translate the individual signs composing a specific contribution into the other respective three sign languages, if this is requested by the user.

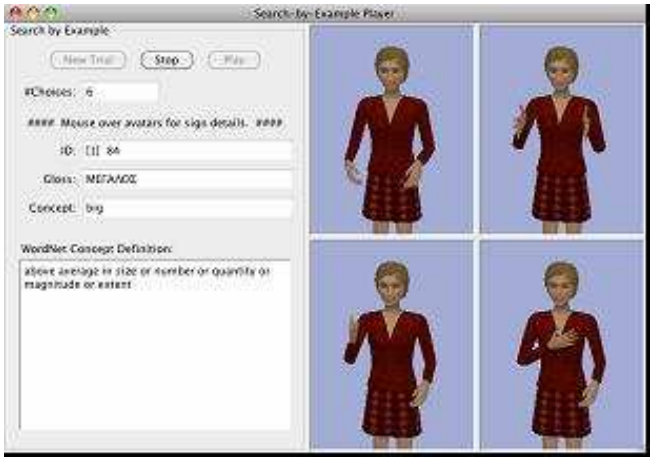


Figure 1: Search-by-Example Tool - Client Display

The project's three proof-of-concept prototypes have successfully demonstrated the feasibility of this initially formulated objective. In all cases, the user may insert his/her own signs or sign utterances either by choosing lemmas from the system's lexicon or by signing in dictation tempo in front of a Kinect camera. All prototypes are Java applications, which use a client-server architecture.
The aim of the *Search-by-Example* prototype (Figure 1), is to allow a Deaf user to perform a sign or a component of a sign, and to search the Dicta-Sign lexical database on the basis of features extracted from the performance. Once candidate lexical signs have been identified, the database delivers information about the form of the sign, in HamNoSys notation, and other linguistic information, that becomes available to the user. An animation of the sign is provided using an avatar.
While there are a few initiatives towards exploitation of sign recognition technologies for lexicon search purposes (Athitsos et al., 2010; Wang et al., 2010), most lexicon applications, known to us, use video to present sign lemmas, while lemma search is based on either groupings according to basic handshape or typing of the gloss or the rough equivalent lemma in the case of bilingual dictionaries. Typical examples are dictionary applications like *The Online Dictionary of New Zealand Sign Language* (http://nzsl.vuw.ac.nz/) and the *English to ASL Dictionary* (http://www.lifeprint.com/dictionary.htm), and also.
Rather than using video, some systems adopt avatar technology. The terminology dictionaries of ASL such as *The Signing Math Dictionary of ASL* (http://www.vcom3d.com/index.php?id=aslani), which are developed by Vcom3D in order to be used in Deaf education, use advanced avatar technology to display lemmas. Other examples of avatar use for lemma representation apply primitive signing avatar versions, such as the one used in *The National Business Aviation Association K8AIT Sign Llanguage Dictionary* (http://wings.avkids.com/Book/Signing/abc.html), which provides mainly aviation related lemmas and has no search facilities other than by groups of letters of the Latin alphabet, while it displays lemma related concept definitions in written English.
The *Sign Look-up Tool* (Figure 2) also has two components – sign input and recognition, and lexical sign data presentation – in a client-server relationship. The most important extension introduced in this system is the ability to handle multiple languages.

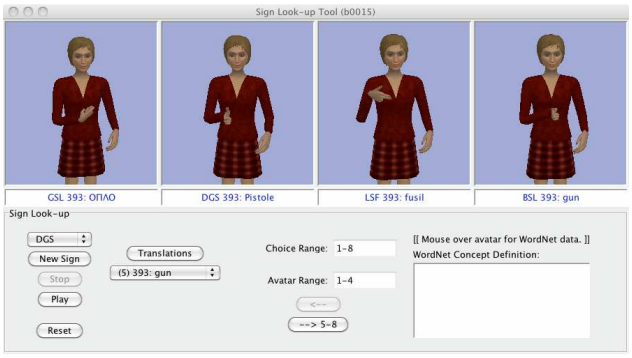


Figure 2: Sign Look-up Tool - Sign Matches Display Mode

Sign Matches Display mode allows the user to view all entries in the ranked list of signs potentially matching the latest user input, based on the list supplied by the server. *Translations Display* mode allows the user to view translations of any of the current set of matching signs in the four sign languages covered by the Dicta-Sign Lexicon. Moreover, the Dicta-Sign Lexicon contains WordNet descriptions for each concept. The user can see the WordNet entry for a concept, in the panel at the bottom right of the window, by hovering over the avatar currently displaying that concept.

Finally, the *Sign-Wiki* prototype (Figure 3) integrates all technologies and resources developed during the project.

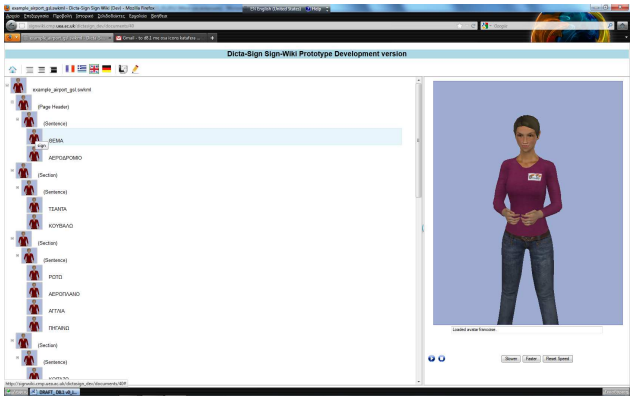
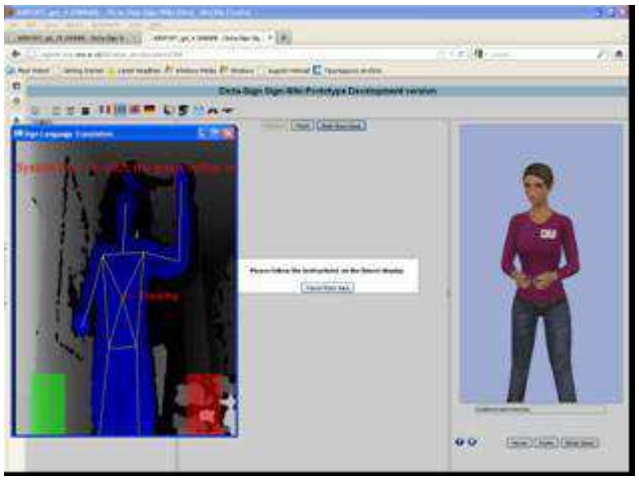


Figure 3: Sign-Wiki typical page

The Dicta-Sign sign-Wiki demonstrates the potential of sign languages to participate in contemporary Web 2.0 applications where user contributions are editable (Figures 4, 5) by an entire community and sign language users can benefit from collaborative editing facilities.
The server developed within the project provides the same service as a traditional Wiki, but using sign language. Sign language content may be directly inserted in dictation style by means of a Kinect camera (Ong 2012). Newly created or previously existing content may be edited, saved and uploaded for presentation. Instead of using text as the output medium, a signing avatar presents information. The system matches the user's signs against a stored dictionary, and the matched signs are used to generate the movements of the signing avatar. The use of an avatar preserves the anonymity of the user and facilitates modification and reuse of information present on the site. If a Kinect device is not available, sequences can be created by selecting signs from the system's lexicon, using spoken language synonyms to expand the range of choices.



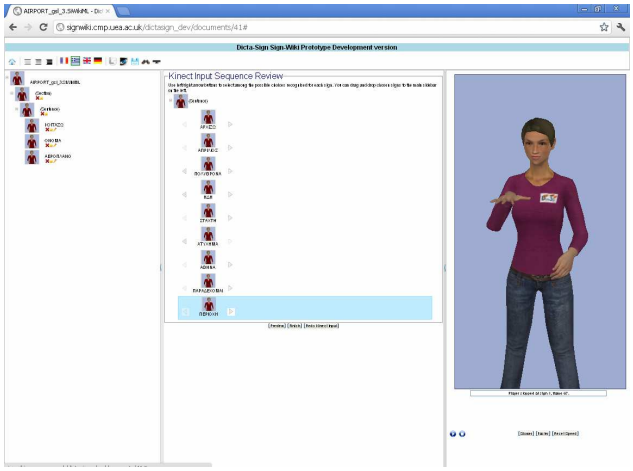


Figure 4: Sign-Wiki - Input reviewing

Innovation in respect to communication via sign language is summarised in the following:

I) Sign-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.

II) In the latter case, the user may find support by a Sign-Look-up translation module, which allows search of signs in four sign languages. Multilingual correspondences of the same concept increase the possibility of its understanding.

III) The 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.

IV) The user may create new SL content by either entering his/her own productions to the system by means of a Kinect devise by exploiting the project's sign language recognition technologies and/or

using the sign creation tools and linguistic models also used for editing purposes. In the case of real time input from the part of the user, single signs or sign phrases are performed in dictation style.

V) The user may save, upload and present his/her content preserving his/her anonymity, since performance of sign language content happens by means of a signing avatar exploiting sign animation technologies.

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.

Figure 5: Sign-Wiki – Sign Builder Tool

Finally, 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 user's preferences in respect to interaction with the systems, thus, gaining advanced human-computer interface design for Web 2.0 sign language applications, that can be best viewed in implementation of the sign-Wiki prototype, also serving as the project demonstrator.

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.

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.

5. Acknowledgements

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.

6. References

Antonakos, E., V. Pitsikalis, I. Rodomagoulakis and P. Maragos (2012). "Unsupervised Classification of Extreme Facial Events using Active Appearance Models Tracking for Sign Language Videos". submitted for publication in: *Proc. Int'l Conf. on Image Processing (ICIP-2012)*, Orlando, Florida, USA .

Athitsos, V., Neidle, C., Sclaroff, S., Nash, J., Stefan, A., Thangali, A., Wang, H., and Yuan, Q. (2010). "Large Lexicon Project: American Sign Language Video Corpus and Sign Language Indexing/Retrieval Algorithms". *Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies (CSLT)*, pages 11-14.

Collet, C., M. Gonzalez and F. Milachon (2010). "Distributed system architecture for assisted annotation of video corpora". *Proc. of the LREC Workshop on the Representation and Processing of Sign Languages : Corpora and Sign Languages Technologies*, May 17th–23rd 2010, Valetta, Malta., pp. 49-52.

Cooper H, Pugeault N, Bowden R (2011). "Reading the Signs: A Video Based Sign Dictionary". In *Proc IEEE Workshop on Analysis and Retrieval of Tracked Events and Motion in Imagery Streams (ARTEMIS)*, Proc ICCV11.

Cooper, H., N. Pugeault, E.-J. Ong and R. Bowden (2012). "Sign Language Recognition using Sub-Units." *Journal of Machine Learning and Recognition: Gesture Recognition*. (to appear).

Dubot, R., Collet, C. (2012). "Improvements of the Distributed architecture for assisted annotation of video corpora". *Proc. of the LREC Workshop on the Representation and Processing of Sign Languages: Interactions between Corpus and Lexicon*. May 21-27 2012, Istanbul, Turkey.

Efthimiou, E., S.-E. Fotinea, T. Hanke, J. Glauert, R. Bowden, A. Braffort, C. Collet, J. Maragos and F. Goudenove (2010). "Dicta-Sign: Sign Language Recognition, Generation, and Modelling: A Research

Effort with Applications in Deaf Communication". *Proceedings of the Language Resources and Evaluation Conference Workshop on the Representation and Processing of Sign Languages : Corpora and Sign Languages Technologies*, May 17th–23rd 2010, Valetta, Malta.

Elliott R, Cooper H M, Ong E J, Glauert J, Bowden R, Lefebvre-Albaret F (2011), Search-By-Example in Multilingual Sign Language Databases, In *Proc. Sign Language Translation and Avatar Technologies Workshops*, 2011.

Filhol, M. (2012, tbp). "Combining two synchronisation methods in a linguistic model to describe Sign Language" In E. Efthimiou, G. Kouroupetroglou, S.-E. Fotinea (Eds.), *Gesture and Sign Language in Human-Computer Interaction and Embodied Communication*, Springer LNCS/LNAI 7206.

Filhol, M. & Braffort, A. (2012, in these proc.). "A Study on Qualification/Naming Structures in Sign Languages". *Proceedings of the Language Resources and Evaluation Conference Workshop on the Representation and Processing of Sign Languages: Interactions between Corpus and Lexicon*, May 27th 2012, Istanbul, Turkey.

Glauert J & Elliott R (2011). "Extending the SiGML Notation - a Progress Report" Proc. *Sign Language Translation and Avatar Technologies*, Dundee, October 2011.

Gonzalez, M. and C. Collet (2011). "Robust body parts tracking using particle filter and dynamic template" *18th IEEE International Conference on Image Processing (ICIP 2011),* 11-14 Septembre 2011, Brussels, Belgium. pp 537-540.

Gonzalez, M. and C. Collet (2012). "Signs segmentation using dynamics and hand configuration for semi-automatic annotation of sign language corpora". *Gesture and Sign Language in Human-Computer Interaction and Embodied Communication, GW 2011*. E. Efthimiou, G. Kouroupetroglou and S.-E. Fotinea. Lecture Notes in Computer Science/Lecture Notes in Artificial Intelligence Vol. 7206: Springer-Verlag (to appear).

Hadfield, S. and R. Bowden (2012). "Go With The Flow: Hand Trajectories in 3D via Clustered Scene Flow". Proc. *International Conference on Image Analysis and Recognition (ICIAR)*, June 25-27, 2012, Aveiro, Portugal (Accepted to Appear).

Hadfield S, Bowden R (2011). "Kinecting the Dots: Particle Based Scene Flow From Depth Sensors". In Proc. *Int. Conference Computer Vision, ICCV2011*, Barcelona, Nov 2011.

Hanke, T. & Storz, J. (2008). "iLex – A Database Tool for Integrating Sign Language Corpus Linguistics and Sign Language Lexicography". In O. Crasborn, T. Hanke, E. Efthimiou, I. Zwitserlood & E. Thoutenhoofd (Eds.), *Construction and Exploitation of Sign Language Corpora. 3rd Workshop on the Representation and Processing of Sign Languages*. ELRA, Paris, pp. 64-67.

Hanke T, Matthes S, Regen A, Storz J, Worseck S, Elliott R, Glauert J, Kennaway R. (2011). "Using Timing Information to Improve the Performance of Avatars" in Proc *Sign Language Translation and Avatar Technologies*, Dundee, October 2011

Holt, B., E.-J. Ong, H. Cooper and R. Bowden (2011). "Putting the pieces together: Connected Poselets for Human Pose Estimation". In Proceedings of the *13th Int. Conference on Computer Vision,: Workshop on Consumer Depth Cameras for Computer Vision*, November 6-11, 2011, Barcelona, Spain.

Jennings, V., Elliott, R., Kennaway, R.and Glauert. J. (2010). "Requirements for a Signing Avatar". Proceedings *Workshop on Corpora and Sign Language Technologies (CSLT)*, LREC 2010, Malta, pp133-136, May 2010.

Maragos, P., Pitsikalis, V., Theodorakis, S., Rodomagoulakis, I., and Antonakos, E. (2012). Report on Integrated Continuous Sign Recognition. DICTA-SIGN Tech. Rep. D2.3, NTUA, Greece, Feb.2012.

Matthes, S., Hanke, T., Storz, J., Efthimiou, E., Dimiou, N., Karioris, P., Braffort, A., Choisier, A., Pelhate, J. and Safar, E. (2010). "Elicitation Tasks and Materials designed for Dicta-Sign's Multi-lingual Corpus". In Dreuw, P., Efthimiou, E., Hanke, T., Johnston, T., Martínez Ruiz, G. and Schembri, A., (eds.), LREC-2010, *Proceedings of 4th Workshop on Representation and Processing of Sign Languages: Corpora and Sign Language Technologies*, pp. 158-163.

Ong, E.-J. and R. Bowden (2011). "Robust Facial Feature Tracking using Shape-Constrained Multi-Resolution Selected Linear Predictors". *IEEE Trans Pattern Analysis and Machine Learning* 33(9): pp1844-1859.

Ong. E.-J. and Cooper, H. M. and Pugeault, N., and Bowden, R., (2012). "Sign Language Recognition using Sequential Pattern Trees". In Proc. *Computer Vision and Pattern Recognition 2012*. To appear.

Pitsikalis, V., Theodorakis, S., Maragos, P. (2010). "Data-driven sub-units and modeling structure for continuous sign language recognition with multiple cues". In: Proc.*LREC Workshop: Corpora and Sign Language Technologies*.

Pitsikalis, V., Theodorakis, S., Vogler, C. and Maragos, P. (2011). "Advances in Phonetics-based Sub-Unit Modeling for Transcription Alignment and Sign Language Recognition". *IEEE CVPR Workshop on Gesture Recognition*, Colorado Springs, USA, 2011.

Pugeault N, Bowden R, (2011). "Spelling It Out: Real-Time ASL Fingerspelling Recognition", In *IEEE Workshop on Consumer Depth Cameras for Computer Vision, Proc ICCV 2011*.

Rodomagoulakis, I., S. Theodorakis, V. Pitsikalis and P. Maragos (2011). "Experiments on Global and Local Active Appearance Models for Analysis of Sign Language Facial Expressions". Proc. *9th International Gesture Workshop (GW 2011): Gesture in Embodied Communication and Human-Computer Interaction*, May 2011, Athens, Greece.

Roussos, A., S. Theodorakis, V. Pitsikalis and P. Maragos (2010). "Affine-Invariant Modeling of Shape-Appearance Images applied on Sign Language Handshape Classification". Proc. *International Conference on Image Processing (ICIP-2010)*, Sep. 2010, Hong Kong.

Roussos, A., S. Theodorakis, V. Pitsikalis and P. Maragos (2010). "Hand Tracking and Affine Shape-Appearance Handshape Sub-Units in Continuous Sign Language Recognition". Proc. *11th European Conference on Computer Vision (ECCV): Workshop on Sign, Gesture and Activity (SGA)*, Sep. 2010, Heraklion, Crete, Greece.

Theodorakis, S., Pitsikalis, V. and Maragos, P. (2011a). "Dynamic vs. Static Sub- Unit Modelling for Continuous Sign Language Recognition with Movement- Position Cues". *IEEE Trans. On Systems Man and Cybernetics,* 2011 (under review).

Theodorakis, S., Pitsikalis, V. and Maragos, P. (2011b). "Advances in Dynamic-Static Integration of Manual Cues for Sign Language Recognition". *Proc. 9th International Gesture Workshop (GW 2011): Gesture in Embodied Communication and Human-Computer Interaction*, May 2011, Athens, Greece.

S. Theodorakis, V. Pitsikalis, I. Rodomagoulakis, and P. Maragos (2012). "Recognition with Raw Canonical Phonetic Movement and Handshape Subunits on Videos of Continuous Sign Language", submitted for publication in: Proc. *Int'l Conf. on Image Processing (ICIP-2012)*, Orlando, Florida, USA, Sep. 2012.

Vogler C, (2010) Extraction of segmental phonetic structures from HamNoSys annotations. DICTA-SIGN Tech. Rep. D4.2, ILSP, Greece.

Wang, H., Stefan, A., Moradi, S., Athitsos, V., Neidle, C. and Kamangar, K. (2010). "A System for Large Vocabulary Sign Search". Proc. *ECCV-2010 Workshop on Sign, Gesture and Activity (SGA)*, Sep. 2010, Heraklion, Crete.

Sign language Dictionaries:

The Online Dictionary of New Zealand Sign Language: http://nzsl.vuw.ac.nz/

English to ASL Dictionary http://www.lifeprint.com/dictionary.htm

The National Business Aviation Association K8AIT sign language dictionary http://wings.avkids.com/Book/Signing/abc.html