ELSEVIER

Educational resources and implementation of a Greek sign language synthesis architecture

K. Karpouzis [a,1], G. Caridakis [a,*], S.-E. Fotinea [b,2], E. Efthimiou [b,2]

[a] *Image, Video and Multimedia Systems Lab, National Technical University of Athens, Iroon Polytechniou 9, GR-157 80 Athens, Greece*
[b] *Institute for Language and Speech Processing, Artemidos 6 & Epidavrou, Athens, Greece*

Abstract

In this paper, we present how creation and dynamic synthesis of linguistic resources of Greek Sign Language (GSL) may serve to support development and provide content to an educational multitask platform for the teaching of GSL in early elementary school classes. The presented system utilizes standard virtual character (VC) animation technologies for the synthesis of sign sequences/streams, exploiting digital linguistic resources of both lexicon and grammar of GSL. Input to the system is written Greek text, which is transformed into GSL and animated on screen. To achieve this, a syntactic parser decodes the structural patterns of written Greek and matches them into equivalent patterns of GSL, which are then signed by a VC. The adopted notation system for the representation of GSL phonology incorporated in the system's lexical knowledge database, is Hamburg Notation System (HamNoSys). For the implementation of the virtual signer tool, the definition of the VC follows the h-anim standard and is implemented in a web browser using a standard VRML plug-in.
© 2005 Elsevier Ltd. All rights reserved.

Keywords: Distance education and telelearning; Human–computer interface; Interactive learning environments; Multimedia/hypermedia systems; Virtual reality

* Corresponding author. Tel.: +30 210 7723037.
  E-mail addresses: kkarpou@cs.ntua.gr (K. Karpouzis), gcari@image.ece.ntua.gr (G. Caridakis), evita@ilsp.gr (S.-E. Fotinea), eleni_e@ilsp.gr (E. Efthimiou).
[1] Tel.: +30 210 7723037.
[2] Tel.: +30 210 6875300.

1. Introduction

Greek Sign Language (GSL) is a natural visual language used by the members of the Greek Deaf Community with several thousands of native or non-native signers. Research on the grammar of GSL per se is limited, whereas some work has already been done on individual aspects of its syntax, as well as on applied and educational linguistics. It is assumed that GSL as we now know it, is a combination of the older type of Greek sign language dialects with French sign language influence. Comparison of core vocabulary lists exhibit many similarities with sign languages of neighboring countries, while in morphosyntax GSL shares the same cross-linguistic tendencies as many other well analyzed sign languages (Bellugi & Fischer, 1972; Liddell, 1980; Stokoe & Kuschel, 1978).

GSL has developed in a social and linguistic context similar to most other sign languages. It is used widely in the Greek Deaf Community and the estimation for GSL natural signers is about 40,600 (1986 survey of Gallaudet University). There is also a large number of hearing non-native signers of GSL, mainly students of GSL and families of deaf people. Although the exact number of hearing students of GSL in Greece is unknown, records of the Hellenic Federation of the Deaf (HFD) show that, in the year 2003 about 300 people were registered for classes of GSL as a second language. The recent increase of deaf students in mainstreamed education, as well as the population of deaf students scattered in other institutions, minor town units for the deaf and private tuition may well double the total number of secondary and potential sign language users in Greece. Official settings, where GSL is being used include 11 Deaf clubs in Greek urban centers and a total of 14 Deaf primary, secondary and tertiary educational settings.

The VRML based sign synthesis system presented here is integrated to an educational platform intended to support young pupils of early primary school acquire the proper linguistic background so that they can take full advantage of language based educational material. The platform offers students the possibility of systematic and structured learning of GSL, for either self-tutoring or participation to virtual classroom sessions of asynchronous teaching, and its design is compatible with the principles that generally define systems of open and distant learning. Besides teaching GSL as a first language, in its present form the platform can be used for the learning of written Greek through GSL, and it will also be open to future applications in areas of other subjects in the school curriculum. Support of teaching written Greek to young deaf pupils is of significant importance given the difficulty of deaf people to make associations between concepts and written forms. This happens because the written form of an utterance is a convention for the representation of sounds, which is incomprehensible in the case, where no perception of sound is possible. According to statistics of the Hellenic Pedagogical Institute (Kourbetis, 1999), the average reading capacity of deaf adults corresponds to mid primary school level.

The virtual signer performs the output of a text-to-signing conversion procedure. Here, we describe the procedures followed during the compilation of the educational material and the implementation of the sign language synthesis component of the educational platform. In this process, we utilized existing software components for the web-based animation of an h-anim virtual character; the adoption of widely accepted character definition and animation standards caters for the extensibility and reusability of the system resources and its content. Fig. 1 describes the abstract architecture and dataflow between the components of the integrated conversion system.

SPOKEN LANGUAGE LAYER

```
┌─────────────────────┐
│  Spoken language    │
│      phrases        │
└─────────────────────┘
          │
          ▼
┌─────────────────────┐
│  Common phrase      │
│     selection       │
└─────────────────────┘
          │
          ▼
┌─────────────────────┐
│  Syntactic and      │
│  lexical analysis   │
└─────────────────────┘
          │
          ▼
┌─────────────────────┐
│  Conversion to Sign │
│  Language syntax    │
└─────────────────────┘
          │
          ▼
┌─────────────────────┐
│   Sign synthesis    │
└─────────────────────┘
```

SIGN LANGUAGE LAYER
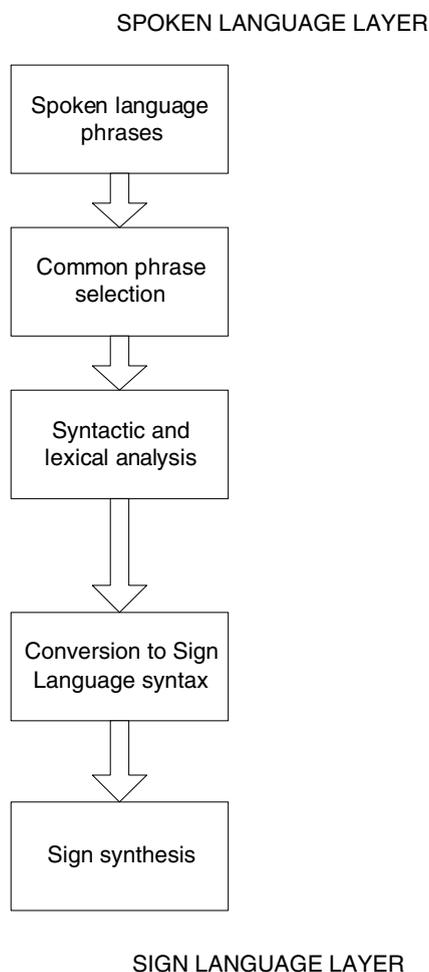
Fig. 1. Overview of the proposed architecture.

2. GSL in the educational context

Formal teaching of GSL as a first language from the very early school years and development of relevant educational content is becoming very urgent since law 2817/2000 was put into action by the Hellenic State. This defines that "the official school language of deaf and hard of hearing students is the Greek Sign Language" and that "knowledge of the Greek Sign Language is a prerequisite for the positioning of tutors and special education staff at the schools that host deaf and hard of hearing students". In this context, the recent education programs of the Hellenic Pedagogical Institute require that all educational material, which will be produced from now on, must be accessible to the deaf students through the use of the Greek Sign Language. Till very recently, however, students with hearing impairments could receive educational material only in written Greek form.

2.1. Linguistic research and background in the area of SL

In Greece, there have been some serious attempts in the domain of SL lexicography in the recent past (NOEMA: a Multimedia Dictionary of GSL Basic Vocabulary, A Children's Dictionary of GSL, Dictionary of Computing Signs), mainly for educational purposes. A complete decoding of GSL structure, is not yet publicly available though.

The linguistic analysis in the framework of the here reported research, is based on overall assumptions for the adequacy of sign languages as by Stokoe (1978) and Kyle and Woll (1985), among many, where GSL is analyzed to its linear and non-linear (simultaneous) components (Sutton-Spence & Woll, 1999; Valli & Lucas, 1995). The linear part of the language involves any sequences of lexical and functional tokens and their syntactic relations, while non-linear structures in GSL, as in all known sign languages, are present in all levels of the grammar. The basic semantic unit of a Sign Language is the sign, a gesture or movement that conveys a concept and functions, much like words in spoken languages. Each sign is described as to its handshape, location, movement, orientation, number of hands and use of any obligatory non-manually articulated elements (e.g., mouth patterns, head and shoulder movements, facial expression and other non-manual features), based on the Stokoe model.

It has been considered essential that the sign synthesis final output is as close to native GSL as used by the Greek Deaf Community. In this respect, forms of 'signed Greek' or other manual codes for the teaching of Greek are excluded and the two languages (GSL and Greek) are treated as first and second language, respectively, for our target user group.

2.2. GSL language resources for educational applications

Implementation of both the tutoring and the summarization tools of the educational platform requires collection of extensive electronic language resources for GSL as regards the lexicon and the structural rules of the language (Sapountzaki, Efthimiou, Karpouzis, & Kourbetis, 2004). The actual data of the study are based on basic research on GSL analysis undertaken since 1999 as well as on experience gained by projects NOEMA and PROKLISI (Efthimiou, Vacalopoulou, Fotinea, & Steinhauer, 2004b). The data consist of digitized language productions of deaf native GSL signers and of the existing databases of bilingual GSL dictionaries, triangulated with the participation of deaf GSL signers in focus group discussions. Development of language resources follows methodological principles on data collection and analysis suitable to the minority status of GSL. Wherever the status of individual GSL signs is in consideration, the Hellenic Federation of the Deaf (HFD) is advised upon, too.

Many of the grammar rules of GSL are derived from the analysis of a digital corpus that has been created by videotaping native signers in a discussion situation or when performing a narration. This procedure is required, because there exists little previous formal analysis of GSL and rule extraction has to be based on actual data productions of native signers. The basic design of the system, except for the educational content this currently supports, focuses on the ability to generate sign phrases, which respect the GSL grammar rules in a degree of accuracy that allows them to be recognized by native signers as correct utterances of the language.

In this respect, implementation of an educational tool proves a challenge for in-depth work on both lexicography and structural analysis of GSL. For the first time, research goes beyond a mere

collection of glosses as is the case with many previous bilingual dictionaries of sign languages (Brien & Brennan, 1992), into the domain of productive lexicon (Wilcox, Scheibmann, Wood, Cokely, & Stokoe, 1994), i.e., the possibility of building new GSL glosses following known structural rules. It also provides a test bed for machine translation (MT) performance in predictable environments, using an effective module for the matching of structural patterns between the written input and the SL-structure output to feed the sign synthesis engine. It is a design prerequisite that the system of GSL description should have an open design, so that it may be easily extendible allowing additions of lemmas and more complicate rules, with the long term objective to create an environment for storage and maintenance of a complete computational grammar of GSL.

2.2.1. Grammar content definition

In the early implementation phase, the subsystem for the teaching of GSL grammar covered a restricted vocabulary and a core grammar capable of analyzing a restricted number of main GSL grammatical phenomena, which might be argued that belong to signing universals. Synthesis of GSL requires the analysis of the GSL signs into their phonological parts and their semantics. It was agreed that only monomorphemic signs that use only one handshape were to be initially analyzed, so that feedback from the technical team would determine further steps. In the second stage, more complicated sequential structures of signs are considered (e.g., compound word signs) and once individual signs are transcribed and stored in a database, it is currently investigated how additional tiers such as basic non-manual features can be added with the least technical difficulty.

At the stage of grammatical analysis, findings from other sign language grammars, as well as the views of our deaf native consultants are taken into account in order to verify findings. It is admitted that there is a huge proportion of work to be done on the pragmatics of GSL, which is noted as one of the future aims of the developers. An example of GSL pragmatics aspect and its relation with real-world situations, regards the use of indexes or classifiers, namely, handshapes with specific semantic value, the use of which in sign formation, differentiates the class of entities indicated by the basic sign. In this sense, signs such as "walk" may be performed either with the classifier [+human] or the classifier [+animal], in order to accurately convey real-word information.

Furthermore, an interesting parameter of a virtual signer is the ability to sign letters of the written alphabet (fingerspelling). This technique is useful in cases of transcription of proper nouns, acronyms, terminology or general terms for which no specific sign exists. Fingerspelling is used extensively in some other sign languages such as the American Sign Language (ASL) or the British Sign Language (BSL); our evidence in GSL suggests that it is only used occasionally, rarely incorporating fingerspelled loans into the core of the language. From a technical point of view, however, it is quite simple for a VC to fingerspell as this process includes no syntax, movement in signing space[3] or non-manual grammatical elements. Many previous attempts of sign animation would go up to the level of fingerspelling or signing only sequential structures of a representation of the written or spoken language. Since then technology has developed and so has linguistic description of sign language structures. On the other hand, few deaf people in Greece use finger-

---

[3] Signing space in Sign Languages is generally defined as a rectangle that starts at waist height, extends to slightly above the head and no more than a foot in front of the body, rarely extending more than one foot left or right beyond the torso (Tennant & Gluszak, 1998).

spelling or a code such as 'Signed Exact Greek' extensively. For these reasons the present work aims to represent a form of GSL as close to natural fluent signing as possible, and only uses finger-spelling occasionally, for example in language games, where teaching of written Greek is the focus.

2.2.2. Notation and glossing

In order to decide on the notation to be followed for sign representation in the lexical resources database, the existing international systems of sign language recording were evaluated. Notation represents a vital part of the whole engine as it serves for the communication between the linguistic subsystem that determines the meaningful movements in the context of GSL and the technological subsystem that performs these movements with a synthetic 3D model signer.

The tool adopted for the transcription and notation of lexical signs is HamNoSys, a pictographic notation system developed by the University of Hamburg for the description of the phonology of signs (Prillwitz, Leven, Zienert, Hanke, & Henning, 1989). This notation forms the corpus of GSL lemmas while for the representation of sequential structures, i.e., in the phrase level, the ELAN language annotator (ELAN URL: http://www.mpi.nl/tools/elan.html) is been adopted. We considered these two systems as most suitable to the text-to-sign animation according to reviews of recent relevant projects. The classic Stokoe model is used for the morphophonological description, with one additional tier with written Greek words of harsh semantic equivalents of utterances. An aim of the developers is to add more tiers as work proceeds, such as those mentioned above on the use of non-manual features and on pragmatics, either using the existing symbols in HamNoSys and ELAN or adding extra symbols to cover specific expression needs. Sign-writing was another transcribing tool under consideration, but was not chosen, given the expected compatibility of HamNoSys within the ELAN tiers in the near future.

3. E-learning in Greece: history and current practice

Early attempts (1994) to exploit on-line connection to sources of educational material in the Greek territory, focused on creation of school networks with the aim to ensure delivery of educational material of equal quality and content volume to pupils at provincial (especially distant and isolated single-tutor institutions) and urban schools. These networks offered students at the province the opportunity to communicate with central schools, receive otherwise inaccessible work and information material and discuss their projects with a wide range of relevant individuals, including tutors and colleagues.

The next generation of on-line educational tools focused on various subjects, language teaching being one of the prevailing areas of application. In this context, a series of innovative products addressed the issues of teaching Greek (Antoniou-Kritikou, Carayannis, & Katsouros, 2001; Charalabopoulou & Konstantakis, 2001) as first or second language, making use of various interactive techniques to present the subject, to test the student's acquisition through related exercise material and score his/her response. In all cases, a central issue remained how to achieve variety of material presentation and ergonomic structuring of exercises, which also signified areas of emphasis in respect to system architecture, parallel to content management subsystem, that allowed the teacher to update educational material and keep a record of the students' skills development.

Currently, the most common type of application involves network connection of schools located at a distance from urban centers, a fact that renders direct contact to extended educational and cultural material extremely difficult, if not possible at all. This type of network makes use of technologies that allow implementation of the above mentioned characteristics, crucially emphasizing at the teacher's ability to update educational content and the student's potential to show and discuss his/her work (Hatzigeorgiu, Carayannis, & Antoniou, 2001). In this context, virtual class session design also exploits video conferencing and Whiteboard options to support interaction.

4. An e-learning platform for GSL

Technological tools as those referred to in the previous section, constitute standard facilities offered by state-of-the-art Internet technologies, to be exploited in current e-learning platforms. They also signal the dominant tendencies in educational practice inside the framework of the gradually establishing Information Society, which is characterized by 'the emergence of an intelligent distributed environment, where access to information in heterogeneous databases and interpersonal communication is concurrently available through a variety of access technologies' (Stephanidis, 2001). In order to convey linguistic information in a sign language we are obliged to make use of 3D representation mechanisms. In the case of an e-learning platform addressing deaf users, educational material has to be delivered in the natural language of the deaf community, in our case GSL. But what are the quality criteria and design principles that can form the basis for development in this case?

4.1. Designing an accessible educational platform

Some of the R&D issues related to the Information Society focus, in particular, on the requirement for Universal Access to Information Society Technologies. In this sense, conventional computer meditated human activities as well as now emerging services and applications such as access to on-line information, e-communication, digital libraries, e-learning, on-line communities, telework, etc. (Stephanidis, 2001) need to be readapted as to become accessible by all possible users.

The requirement for development of Information Society Technology products and services that are accessible to all citizens, involves architectures based on systematic effort to implement methodologies, design principles and tools in a proactive way (instead of a posteriori adaptations). Such a priori approaches define the concept of Design for All in the era of Information Society. In Greece, this need is recognized in research initiatives which focus on exploitation of Design for All and Universal Access principles (Stephanidis & Savidis, 2001).

In our specific context, reference to "all users" translates as inclusion of the population of deaf and hearing impaired students. Consequently, designing an accessible educational platform requires integration of mechanisms that allow access to its content and tools by conveying meaning with 3D representation mechanisms. Our case study, involves rearchitecturing that allows a virtual signer (Efthimiou, Sapountzaki, Karpouzis, & Fotinea, 2004a) to perform as mediator between tutor and student to present educational content in GSL. The selection of the Greek Deaf Community as our target group, was supported on one hand by the high distribution of the

group members to mainstream schools all over Greece. On the other hand, GSL still lacks substantial educational material (Efthimiou & Fotinea, 2004) even though it is the official language of the Greek deaf students by law since 2000 (Act 2817/2000), and consequently should be their primary means for education. Moreover, this action is in accordance with EU principles for Accessibility to Information in Special Education (see, also COM (2000) 284 final, 2000) and the e-Europe initiative.

4.2. Tutoring system design and evaluation principles

In a scenario of Universal Access to e-education, which also incorporates user–system interaction by means of signing into GSL, our current work addresses to the population of deaf Greek pupils of primary schools, aiming at developing a platform with animated signing in GSL by a virtual human tutor (avatar, VC signer) (Huang, Eliens, & Visser, 2002; Kennaway, 2001).

The test bed learning procedure concerns teaching of GSL grammar to early primary school pupils, whereas the platform also incorporates a subsystem that allows approach by the deaf learner to material available only in written Greek form by means of a signed summary. The learning process, in practice, involves an initiator of the session, the students in groups or alone and a teacher–facilitator of the process, physically present with the students. The process can take place in real time or can be relayed. There is provision of a virtual Whiteboard, icon banks and chat board visible in the screen along with the virtual signer for common use in the classroom. The participants are also to see each other in real time through a web camera, in order to verify results of GSL learning.

Specifications for the formation of GSL resources of the application are crucially based on exhaustive research in the official, recently reformed, guidelines for the teaching of Greek language and of GSL in primary schools for the deaf. The educational content of the platform follows the same guidelines as the hearing children's curriculum, so that the same grammatical and semantic units can be taught in the two languages, GSL and spoken Greek. Concepts such as subject–object relations, types of verbs and discourse functions of the language form the units of the curriculum, so that the same principles are taught under the same platform, but without projecting a mirror image of the Greek grammar onto GSL. For the selection and arrangement of the educational material, design follows close cooperation with the Hellenic Pedagogical Institute in Athens, which is the main official agency in charge of the development of educational material.

Evaluation of the platform will be carried out by user groups of native signers, GSL students and their tutors in real use environment. The Hellenic Federation of the Deaf is involved in distributing the prototype to special education schools and supervising students' performance. The platform will also be introduced in mainstream schools for evaluation under conditions of mixed class audience, a procedure supervised by the Hellenic Pedagogical Institute.

4.3. Educational function

Typical on-line language exercise models, included in former educational software, were restricted by technological limitations to text based types mostly (e.g., hangman, 'fill in the blank'

type, multiple choice exercises, etc.), which allowed for easy and standard scoring techniques. In the case of sign language teaching, off-line environments included exercises heavily based on the use of still image or video.

The conventional means for the teaching of GSL include, and are certainly not limited to:

- text;
- pictures/still image;
- pre-recorded video;
- pantomime as well as;
- signing of oral Greek utterances;
- written word transcription by finger alphabet.

The current state-of-the-art, allows expansion of the above to include:

- live video (teleconferencing facilities);
- anthropoid tutor, exploiting avatar technologies to dynamically present educational content in GSL.

Combination of any of the above means, produces classes of exercises that can be dynamically reviewed and serve different educational purposes.

When addressing the issue of educational scenarios in an e-learning platform for GLS, standard approaches comprise the use of a range of exercises to cover for passive as well as active participation in the class. Passive participation scores comprehension declared through indication of acquired objects by the students. Active class participation involves student's response that generates new language items. This is the point of focus in respect to GSL adaptation to the e-learning platform under consideration (Efthimiou & Fotinea, 2004). Towards this end, implementation involves exploitation of the technological tools described above.

In the virtual class environment, interaction based tutoring and related exercises make extensive use of video conferencing and sharing of the Whiteboard. When a mainstream school tutor with no knowledge of GSL has to provide the virtual class with some type of information (descriptions, instructions, etc.), the chat tool is first employed, to feed the sign generator that will convey the required meaning to the deaf audience.

Exercise models that are constructed for use in virtual class environment may be grouped according to content (restricted–unrestricted), type of activity (single user–user group), type of structure (game–task), difficulty, etc.

Real time scoring by the tutor allows for modification of the lecture according to class member needs, during session. This function is also connected with exercises that can dynamically change content.

Drills of exercises entail self assessment when the student needs to correctly complete a task, before being allowed to proceed to a next stage. An automatic corrective procedure employs display of the task and of the student's response by the virtual signer.

Exercises for individuals or groups connected to the virtual class, may be presented in the form of a "contest" among knots to raise interaction and active participation to the educational process.

In our case, the decision to explore the use of VC signer was based on the need to provide educational material, freed from the static nature of pre-recorded videos that can convey only frozen linguistic information. A sample description of elementary exercises for GSL acquisition organized in 3 groups and currently available in the platform prototype, is next presented with the aim to stress the advantages of dynamic sign synthesis, in what concerns modification of existing or addition of new educational material by the tutor.

One group of exercises deals with signs that use the same handshape but start from different positions with respect to the signer's body or the neutral signing space and consist of different movements. An example of such a group in GSL includes the words 'table', 'house', donkey', 'slipper' and 'tent'. In this framework, young pupils are initially presented with the VC signing each word in a particular group and a sketch depicting the signed concept; the use of sketches instead of photographs or written words is adopted since very young deaf pupils have not yet developed any skills related with spoken or written languages and thus, their mother tongue is the relevant sign language; besides a sketch provides a better representation of an abstract concept, than any photograph of a specific instantiation of it. After initial presentation of tutoring object, the pupils go through a number of drills, similar to the ones found in usual language teaching classes. These drills mainly consist of choosing the correct sketch relating to a random sign performed by the VC and matching different instances of the VC with the correct sketch, by picking from an on-screen sketch pool.

A second group of exercises includes signs with similar or semantically related meaning, signed with the same or different handshapes. An example is the group 'human', 'tall', 'fat', 'child', 'female'. The drills here are the same with the ones in the previous exercise group, as is also the case with the third group of demo exercises. In this last category, sign pairs are formed, consisting of signs composed of same phonological features (handshape, movement, location, palm orientation) but differing in their grammatical classification, e.g., 'sit–chair', 'eat–food' and 'love (verb)–love (noun)' by means of movement repetition.

5. Technical considerations

The implementation team has reviewed currently available VC and animation technologies for the representation of sign language in order to adopt one of the most prominent technological solutions. The movements of a synthetic 3D signing model have to be recorded in a higher and reusable level of description, before they are transformed in parameters of body movement. In the area of text-to-sign animation there have been some similar projects (VISICAST, Thetos, SignSynth and VSIGN among them) that we used as background.

5.1. The h-anim standard

The increase in computing power and networking capabilities has brought renewed interest in 3D graphics over the past years, resulting in a steady emergence of applications related to the modeling and animation of 3D human figures. A usual setback that these applications face has to do with data exchange and reuse. The absence of a standardized skeletal representation system usually results in the development of proprietary solutions to help smooth transitions between existing systems and software.

The h-anim ISO standard (ISO/IEC 19774) provides a systematic approach to representing humanoid models in a 3D graphics and multimedia environment. In this framework, each humanoid is abstractly modeled in terms of structure as an articulated character, which can be embedded in different representation systems and animated using the facilities provided by that representation system. Building on this, the h-anim standard defines animation as a functional behavior of time-based, interactive 3D, multimedia formally structured characters, leaving the particular geometry definition in the hands of the modeler/animator. A number of general, high-level goals guide the specification of the h-anim standard, namely compatibility, resulting in the features of an h-anim figure being implementable in any compliant browser or application, flexibility, separating the application semantics from the character definition and simplicity, which enables the standard to be minimal in its provisions, while also extensible in the future. To realize these goals, the specification allows direct access to the joint hierarchy of the human skeleton, as well as the actual geometry that makes up the individual body segments in a way separating animation definition from the model itself.

5.1.1. Definition and modeling of an h-anim model

The skeletal description of an h-anim model consists of a tree of Joint objects, corresponding to the joints found on a human body, that define the transformations from the root joint (called HumanoidRoot) to the end effector of each limb of the hierarchy. In general, only the Humanoid-Root joint is required, while all other Joint objects are optional; however the more Joint objects found in a model definition, the more flexible the model becomes to animate. As a result, the standard defines particular predefined sets of joints as levels of articulation (LOA): a set with fourteen joints is defined as a "low level of articulation", whereas a humanoid model with 72 joints is characterized as a "high level of articulation"; all compliant figures are animated by applying transformations at those joints.

Regarding the modeling process, h-anim humanoids are built with actual human size ranges in mind (roughly 1.75 m tall) and stand facing in the $+Z$ direction, with $+Y$ up and $+X$ to the humanoid's left. The origin of the coordinates system is located at ground level, between the humanoid's feet. In the default stance, arms are straight and parallel to the sides of the body with the palms of the hands facing inwards, while the face is modeled with the eyebrows at rest, the mouth closed and the eyes wide open. For the recording and definition of handshape and gestures, motion tracking and haptic devices (such as CyberGrasp or Acceleration Sensing Glove with a virtual keyboard) were initially considered; however, it was agreed that, if the HamNoSys notation commands would provide acceptable quality, based on the initial implementation, motion capture sequences will not need to be applied. In any case, semantic notation is a far more flexible and reusable solution than video files or motion capture, since an h-anim VC can take advantage of the dynamic nature of phonological and syntactic rules.

5.2. Implementation

For the content designer to interact with a VC, a scripting language is required. In our implementation, we chose the STEP (Scripting Technology for Embodied Persona) language (Huang et al., 2002) as the intermediate level between the end user and the virtual actor. A major advantage of scripting languages such as STEP is that one can separate the description of the individual

gestures and signs from the definition of the geometry and hierarchy of the VC; as a result, one may alter the definition of any action, without the need to re-model the virtual actor. The VC utilized here is compliant with the h-anim standard, so one can use any of the readily available or model a new one.

Scripted animation is an interchangeable and extensible alternative of animation based on motion capture techniques. One can think of the relation between these two approaches similarly to the one between synthetic animation and video-based instructions: motion capture can be extremely detailed with respect to the amount and depth of information, but is difficult to adjust or adapt when produced and typically requires huge amounts of storage space and transmission capacity to deliver. On the other hand, scripted animation usually requires manual intervention to compile and thus is minimal and abstract in the way it represents the various actions of the avatar. As a result, such scripts require a few hundred characters to describe and can be reused to produce different instances of similar shape (Foulds, 2004). This is illustrated in the code snippet in Fig. 2, which illustrates the required transformations for the right-hand to assume the 'd'-handshape. As is easily demonstrated, the same code of the left hand can be compiled by mirroring the described motion, while other, more complicated handshapes can start with this representation and merely introduce the extra components into it.

In the reported project, a syntactic parser (Boutsis, Prokopidis, Giouli, & Piperidis, 2000) decodes the structural patterns of written Greek and matches them into their equivalents in GSL. These are fed into an automated system that decodes HamNoSys notation sequences for each lemma; this system essentially transforms single or combined HamNoSys symbols to sequences of scripted commands. A typical HamNoSys notation sequence consists of symbols describing the starting point configuration of a sign and the action that the signing consists of. Symbols describing the initial configuration refer to the handshape that is used during the sign and the starting position and orientation of the hand that performs the sign; if the other hand takes part in the sign, as is the case in the GSL version of 'doctor', it is the relative position of the two hands that matters, for example 'the main hand touches the elbow of the secondary arm'. Other information includes symmetry, if both hands follow the same movement pattern and any non-manual components. Fig. 3 shows a frame of the signing sequence for 'donkey'; the VC shown here is 'yt', by Matthew T. Beitler, available at http://www.cis.upenn.edu/~beitler. A demonstration with limited vocabulary and some phrase examples can be found online at http://www.image.ece.ntua.gr/~gcari/gslv.

Fig. 4 shows the HamNoSys sequence for the particular sign, shown on the top of the page of the user interface. The first symbol here indicates that both hands perform the same movement, starting from symmetrical initial locations with respect to the signer's torso. The second symbol indicates the handshape, which here is an open palm, referred to as the 'd'-handshape in GSL, while the next shows palm orientation. The following symbols handle the starting position of the palm, which here almost touches the temple of the signer's head. Symbols contained in parentheses describe composite movements, while the last character forces the signer to repeat the described movement.

Fig. 5 shows the VC signing the GSL version of 'child', (see Fig. 6) while Fig. 7 shows an instance for the sign for 'children'. The design of the automated script production system enables us to use the description of the former sign (Fig. 6) to construct the definition of its plural form. In this case, the plural form is shown by repeating the same downward hand movement, while

```
par([

    turn(humanoid,r_thumb1,rotation(1.9,1,1.4,0.6),very_fast),

    turn(humanoid,r_thumb2,rotation(1,0.4,2.2,0.8),very_fast),

    turn(humanoid,r_thumb3,rotation(1.4,0,0.2,0.4),very_fast),

    turn(humanoid,r_index1, rotation(0,0,0,0),very_fast),

    turn(humanoid,r_index2,rotation(0,0,0,0),very_fast),

    turn(humanoid,r_index3,rotation(0,0,0,0),very_fast),

    turn(humanoid,r_middle1,rotation(0,0,1,1.5999),very_fast),

    turn(humanoid,r_middle2,rotation(0,0,1,1.5999),very_fast),

    turn(humanoid,r_middle3,rotation(0,0,1,1.5999),very_fast),

    turn(humanoid,r_ring1,rotation(0,0,1,1.7999),very_fast),

    turn(humanoid,r_ring2,rotation(0,0,1,1.5999),very_fast),

    turn(humanoid,r_ring3,rotation(0,0,1,0.6000),very_fast),

    turn(humanoid,r_pinky1,rotation(0,0,1,1.9998),very_fast),

    turn(humanoid,r_pinky2,rotation(0,0,1,1.5999),very_fast),

    turn(humanoid,r_pinky3,rotation(0,0,1,0.7998),very_fast)

])
```

Fig. 2. STEP code for a handshape.

moving the hand slightly to the signer's right; direction is indicated by the symbol preceding the parenthesis, while its content describes this secondary movement. As a result, it is only necessary for the parser to indicate the particular modification of the initial sign required to produce the plural form of the lemma. In GSL, these forms are limited, thus enabling us to come up with efficient production rules, such as the one described above. Another possibility is to change the handshape for a sign, especially when the signer wants to indicate a particular quantity or number. Fig. 8 shows the VC signing the GSL version of 'day', while Fig. 9 shows the GSL version of 'two days'; the difference here is that in the latter case the VC uses a two-finger handshape, instead of the straight-index finger handshape, to perform the same movement, starting from the same initial

Greek Sign Language Visualization

Fig. 3. An instance of 'yt' signing 'donkey'.

Fig. 4. The HamNoSys sequence for the GSL version for 'donkey'.

Greek Sign Language Visualization

Fig. 5. The GSL version of 'child'.

position. This difference is more evident in Fig. 10, which shows the VC in a frontal view; this is actually an interesting feature of the Blaxxun Contact 5 (Blaxxun, 2004), the VRML plug-in shown in these figures. Despite the default tilted view being the one of choice from the part of the users, the ability to show frontal and side view of the sign is crucial in learning environments, since

```
par([

turn(humanoid,r_elbow,rotation(-0.722,0.2565,0.1206,1.5760),fast),

turn(humanoid,r_shoulder,rotation(-0.722,0.2565,0.1206,0.0477),fast),

turn(humanoid,r_wrist,rotation(0,1.5,-1,1.570),fast)

]),

sleep(500),

par([

turn(humanoid,r_shoulder,rotation(-0.598,0.2199,0.1743,0.0812),fast),

turn(humanoid,r_elbow,rotation(-0.598,0.2199,0.1743,1.2092),fast)

])
```

Fig. 6. The STEP code for the sign 'child'.

Fig. 7. The GSL version of 'children'.

it caters for displaying the differences between similar signs and brings out the spatial characteristics of the sign (Kennaway, 2001, 2003).

6. Implications and extensibility of the educational platform

As an educational tool above all, the here reported platform intends to offer a user-friendly environment for young deaf pupils aged 6–9, so they can have visual translation of written words

Greek Sign Language Visualization
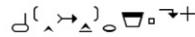
Fig. 8. The GSL version of 'day'.

**Greek Sign Language Visualization**
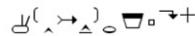
Fig. 9. The GSL version of 'two days'.

and phrases. For deaf young students as a group with special needs, the platform draws some of the accessibility barriers, and the possibility of home use even makes it accessible to family, thus encouraging communication in GSL, but also access to the majority (Greek) language. New written texts can be launched, so the platform may receive in principle unlimited educational content besides primary school grammar units. On the other hand, unlimited school units, such as the

Greek Sign Language Visualization

Fig. 10. The frontal view of the GSL version of 'two days'.

increasing special units with individual deaf students in remote areas can link with one another via the platform.

Moreover, text-to-sign translation can be extended and applied to different environments such as Greek language teaching to deaf students of higher grades, GSL teaching for hearing students, Greek for specific purposes such as to adult literacy classes for the Deaf, etc. In this context, more domains of GSL grammar can be described and decoded, making the output closer to natural signed utterances as our analysis proceeds. This is a challenge not only for theoretical research, but also for computer science and applied linguistic research.

Furthermore, a database with the bulk of GSL utterances, described as to their features from the phonological up to the pragmatic level will be a major outcome of the whole enterprise. Structuring resources of this kind of knowledge will allow matching between representations of GSL structures and their equivalent ones of written Greek. It will also be a challenge to compare directly the grammars of the two languages. In much the same way structures of GSL will easily be compared with counterparts from ASL or BSL (Brien & Brennan, 1992) for research across signed languages.

7. Problems and limitations

The main limitations of the study are described below. These are divided into linguistic, educational and technical ones. Most of the limitations are typical to sign animation attempts, and they were expected before the beginning of the project.

Regarding the linguistic and educational aspects of the project, one of the major issues that needs to be addressed is the fact that in some areas of the language there are no standardized signs, so there may be some theoretical objections as to the use of particular entries. However, a platform such as the one described, allows for multiple translations and does not have any limitations

as to the size of files, which was the case, for example in previous DVD-based GSL dictionaries, containing video entries. Moreover, the platform will be open to updates through the script authoring process.

Another issue is the choice of entries to be included in each stage of the platform development depending on the complexity of their phonological characteristics. As mentioned already in the section on grammar content definition, monomorphemic entries were agreed to be included in the first stage. In the next stages there is gradual provision for polymorphemic signs, compound signs, functional morphemes, syntactic use of non-manual elements, sequential and lastly simultaneous constructions of separate lexical signs, each stage to correspond with the level of linguistic research in GSL.

Besides this, the data available in GSL, when compared with data from written Greek, for example, are dauntingly scarce. Error correction mechanisms were sought after in order to assure reliability of results. Such back-up mechanisms are the use of approved dictionaries, the consultancy of Hellenic Pedagogical Institute, which is the official organization that validates all educational programs of primary and secondary education in Greece, and the feedback from the Deaf Community, along with the continuing data from GSL linguistic research.

The most important technical problems include a solution for smooth transition between concurrent signs and fusion between handshapes so that neighboring signs in a sentence appear as naturally articulated as possible. In the context of the reported project, this issue has been tackled using a nice feature of the STEP engine, which at any time can return the setup of the kinematic chain for each arm. As a result, when the sign that is next in a sequence begins, the kinematic chain is transformed to the required position without having to take into account its setup in the final position of the previous sign. In general, this would be problematic in general purpose animation, since the h-anim standard itself does not impose any kinematic constraints; thus, random motion might result in physiologically impossible, puppet-like animation. In the case of signing though, almost all action takes place in the signing space in front of the signer and starting from the head down to the abdomen; in this context, there are no abrupt changes in the chain setup.

Another issue regarding the animation representation has to do with circular or wavy movement. Since the description follows the same concepts as keyframed motion, circular movement or, generally, paths following a curve must be approximated with discrete key positions. This often results in losing the relative position of the hands, as shown in Fig. 11, which depicts the final position for the sign 'boat'; this sign should normally end up with palms touching, but since this process is designed with the position of the palm in mind, keeping hands together is not a straightforward task.

In addition to this, a major factor in sign synthesis is the grammatical use of non-manual signs, such as meaningful or spontaneous facial expression (Karpouzis, Raouzaiou, & Kollias, 2003) and eye gaze, particularly when eye gaze has to follow the track of hand movements. Similar problems are anticipated on mouth movements on prosodic features of sign phonology. Mouthing the visible part of spoken Greek words will not be an issue for the project yet, but this too is anticipated as a problem to deal with in the future, given that all of the above non-manual features have various functions of internalized parts of GSL grammar. At the moment, the only possible non-manual sign components possible to animate with the STEP platform are gazing towards the signer's moving hands and forward torso leaning, in the case of asking a question. In general, the STEP engine does not yet feature facial animation, so the project team is considering moving to a

Greek Sign Language Visualization

Fig. 11. Problems in the GSL version of 'boat'.

pure MPEG-4 (MPEG-4, 2004) based platform. A nice example of maturing MPEG-4 synthetic technology is the VC named 'Greta' (De Rosis, Pelachaud, Poggi, Carofiglio, & De Carolis, 2003) which supports all required manual and non-manual components, including visemes, the visual counterpart of phonemes used for lip-reading, high-level facial expression, e.g., 'surprise' associated with an exclamation mark or simple facial and head movement, such as raising the eyebrows or tilting the head upwards to indicate negation (see Fig. 12).

The ultimate challenge, as in all similar projects, remains the automatic translation of the language. It is still too difficult to produce acceptable sentences in the automatic translation of any language at the moment, even more so a minor, less researched language with no written tradition such as GSL. Realistically the development teams of the project can expect as an optimum result the successful use of automatic translation mechanisms in GSL only in a restricted, sub-language oriented environment with predetermined semantic and syntactic characteristics.
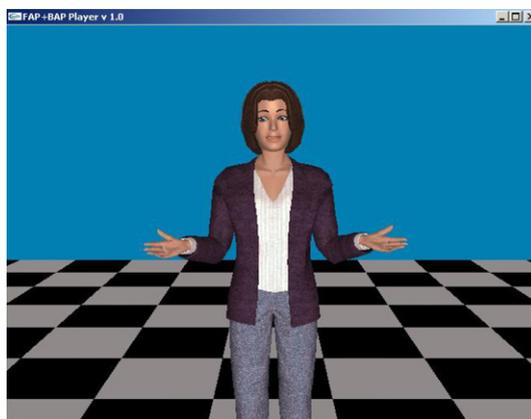
Fig. 12. Greta displaying manual and non-manual signs.

8. Conclusion

In this paper, we have described the underlying design principles and implementation of a web-based virtual signer software component, utilizing language resources suitable for young pupils. This component uses standard linguistic and virtual character technologies to provide semantic and syntactic information from written text and encode it with reusable and extensible sign notation representations. These representations are readable by the VC platform, making them suitable for teaching GSL and providing signed summaries of documents.

Acknowledgements

References

Antoniou-Kritikou, I., Carayannis, G., & Katsouros, V. (2001). A multimedia storyboard as an object and as a starting point for language learning. In *Proceedings of the Eden 10th anniversary conference* (pp. 313–318). Stockholm, Sweden.

Bellugi, U., & Fischer, S. (1972). A comparison of Sign language and spoken language. *Cognition, 1*, 173–200.

Blaxxun Contact 5. (2004). Available from http://www.blaxxun.com/en/products/contact/.

Boutsis, S., Prokopidis, P., Giouli, V., & Piperidis, S. (2000). Robust parser for unrestricted Greek text. In *Proceedings of the 2nd language resources and evaluation conference* (pp. 467–474).

Brien, D., & Brennan, M. (1992). *Dictionary of British sign language/English*. Boston: Faber and Faber.

Charalabopoulou, F., & Konstantakis, N. (2001). Developing a multilingual platform for language learning using the New Technologies. In *Proceedings of 15th symposium of theoretical and applied linguistics, University of Salonica, 4–7 May 2001*.

COM. (2000). 284 final, (2000). Towards a Barrier Free Europe for People with Disabilities, Communication from the Commission to the Council, the European Parliament, the Economic and Social Committee and the Committee of the Regions, Brussels.

De Rosis, F., Pelachaud, C., Poggi, I., Carofiglio, V., & De Carolis, B. (2003). From Greta's mind to her face: modeling the dynamics of affective states in a conversational embodied agent. *International Journal of Human–Computer Studies, 59*(1–2), 81–118.

Efthimiou, E., & Fotinea, S.-E. (2004). An adaptation-based technique on current educational tools to support universal access: the case of a GSL e-Learning platform. In *Proceedings of the TeL'04 Workshop on "Technology Enhanced Learning", 22 August 2004, Toulouse, France, Workshop to WCC-2004 (World Computer Congress)* (pp. 177–186).

Efthimiou, E., Sapountzaki, G., Karpouzis, C., & Fotinea, S.-E. (2004a). Developing an e-Learning platform for the Greek Sign Language Lecture Notes in Computer Science (LNCS) (Vol 3118). Springer pp. 1107–1113.

Efthimiou, E., Vacalopoulou, A., Fotinea, S.-E., & Steinhauer, G. (2004b). A multipurpose design and creation of GSL dictionaries. In *Proceedings of the workshop on the representation and processing of sign languages "from signwriting to image processing. Information techniques and their implications for teaching, documentation and communication", satellite workshop to LREC-2004 conference, 30 May 2004* (pp. 51–58). Lisbon, Portugal.

ELAN annotator, Max Planck Institute for Psycholinguistics. Available from http://www.mpi.nl/tools/elan.html.

Foulds, R. (2004). Biomechanical and perceptual constraints on the bandwidth requirements of sign language. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 65–72.

Hatzigeorgiu, N., Carayannis, G., & Antoniou, I. (2001). Design and implementation of mini-webs for educational uses. In *Proceedings of 2001 symposium on applications and the Internet workshops* (pp. 33–38). San Diego, California.

Huang, Z., Eliens, A., & Visser, C. (2002). STEP: a scripting language for embodied agents. In *Proceedings of the workshop on lifelike animated agents*.

ISO/IEC 19774 – Humanoid Animation (H-Anim). Available from http://www.h-anim.org.

Karpouzis, K., Raouzaiou, A., & Kollias, S. (2003).'Moving' avatars: emotion synthesis in virtual worlds Human–computer interaction: theory and practice(Vol. 2). Lawrence Erlbaum Associates, Inc. pp. 503–507.

Kennaway, R. (2001). Synthetic animation of deaf signing gestures. In *Proceedings of the international gesture workshop*. London: City University.

Kennaway, R. (2003). Experience with, and requirements for, a gesture description language for synthetic animation. In *Proceedings of the 5th international workshop on gesture and sign language based human–computer interaction*. Genova.

Kourbetis, V. (1999). *Noima Stin Ekpaidefsi*. Athens: Hellenic Pedagogical Institute (in Greek).

Kyle, J. G., & Woll, B. (1985). *Sign language: the study of deaf people and their language*. Cambridge University Press.

Liddell, S. (1980). *American sign language syntax*. The Hague: Mouton.

MPEG-4 Home page. Available from http://www.chiariglione.org/mpeg/standards/mpeg-4/mpeg-4.htm.

Prillwitz, S., Leven, R., Zienert, H., Hanke, T., & Henning, J. (1989). HamNoSys. Version 2.0. Hamburg notation system for sign language. An introductory guide.

Sapountzaki, G., Efthimiou, E., Karpouzis, C., & Kourbetis, V. (2004). Open-ended resources in greek sign language: development of an e-learning platform. In *Proceedings of the workshop on the representation and processing of sign languages "from signwriting to image processing. information techniques and their implications for teaching, documentation and communication", satellite workshop to LREC-2004 conference, 30 May 2004, Lisbon, Portugal* (pp. 13–19).

SignSynth. Available from http://www.unm.edu/~grvsmth/signsynth/.

Stephanidis, C. (2001). Editorial. *International Journal of Universal Access in the Information Society, 1*, 1–3.

Stephanidis, C., & Savidis, A. (2001). Universal access in the information society: methods, tools, and interaction technologies. *International Journal of Universal Access in the Information Society, 1*, 40–55.

Stokoe, W. (1978). *Sign language structure* (Revised Ed.). Silver Spring, MD: Linstok.

Stokoe, W., & Kuschel, R. (1978). *For sign language research*. Linstock Press.

Sutton-Spence, R., & Woll, B. (1999). *The linguistics of British sign language: an introduction*. Cambridge University Press.

Tennant, R. A., & Gluszak Brown, M. (1998). *The American sign language handshape dictionary*. Gallaudet University.

Thetos. Available from http://sun.iinf.polsl.gliwice.pl/sign/.

Valli, C., & Lucas, C. (1995). *Linguistics of American sign language* (2nd ed.). Washington, DC: Gallaudet University Press.

VISICAST. Available from http://www.visicast.co.uk/.

VSIGN. Available from http://www.vsign.nl/EN/vsignEN.htm.

Wilcox, S., Scheibmann, J., Wood, D., Cokely, D., & Stokoe, W. (1994). Multimedia dictionary of American sign language. In *Proceedings of ASSETS conference* (pp. 9–16).