Providing Signed Content on the Internet by Synthesized Animation

J. R. KENNAWAY and J. R. W. GLAUERT
University of East Anglia, UK
and
I. ZWITSERLOOD
Viataal, NL

Written information is often of limited accessibility to deaf people who use sign language. The eSign project was undertaken as a response to the need for technologies enabling efficient production and distribution over the Internet of sign language content. By using an avatar-independent scripting notation for signing gestures and a client-side web browser plug-in to translate this notation into motion data for an avatar, we achieve highly efficient delivery of signing, while avoiding the inflexibility of video or motion capture. Tests with members of the deaf community have indicated that the method can provide an acceptable quality of signing.

Categories and Subject Descriptors: I.3.7 [**Computer Graphics**]: Three-Dimensional Graphics and Realism—*Animation*

General Terms: Human Factors, Languages

Additional Key Words and Phrases: Animation, avatar, deaf accessibility, HamNoSys, scripting, SiGML, sign language, virtual reality

ACM Reference Format:

Kennaway, J. R., Glauert, J. R. W., and Zwitserlood, I. 2007. Providing signed content on the internet by synthesized animation. ACM Trans. Comput.-Hum. Interact. 14, 3, Article 15 (September 2007), 29 pages. DOI = 10.1145/1279700.1279705 http://doi.acm.org/10.1145/1279700.1279705

15

1. INTRODUCTION

The aim of the work reported here is to provide signed content for Internet applications, especially Web pages, to facilitate access for deaf people for whom signing is their first language.

We acknowledge funding from the European Union for our work on the eSign project under the Framework V programme.
Authors' addresses: J. R. Kennaway and J. R. W. Glauert, School of Computing Sciences, University of East Anglia, Norwich, UK; email: jrk@cmp.uea.ac.uk, http://www.cmp.uea.ac.uk; I. Zwitserlood, Stichting Viataal, Sint Michielsgestel, The Netherlands; http://www.viataal.nl.

1.1 Needs of Deaf People

Since Internet information is mostly presented visually, it may seem that access for deaf people is not an issue. Indeed, those who have lost hearing later in life can easily access textual information. However, for prelingually deaf individuals (those who were deaf since before learning any language), written material is often much less accessible than if presented in signing. Several studies have shown that the reading performance of deaf children is poor compared to that of their hearing peers [Conrad 1979; Allen 1986; Holt et al. 1996]. From recent reading comprehension testing of deaf pupils aged 7–20 in the Netherlands [Wauters 2005], the mean score is at the level of seven-year-old hearing children. Only 25% of the deaf pupils read at or above the level of a nine-year-old hearing child. Thus situations in which information is presented primarily in written form place them at a substantial disadvantage.

Further, for this community, signing is their language of choice, and reading text in a spoken language is akin to using a foreign language. Taking an example from spoken languages in the UK, there are few Welsh speakers who are unable to read English, but respect for that unique linguistic community means that there is a cultural, and indeed legal, requirement to provide information in the Welsh language.

There are estimated to be 100,000 to 150,000 deaf people in the US for whom sign language is their first language, 40,000 in the UK, 50,000 in Germany, and 17,500 in the Netherlands [Deuchar 1984; Gordon Jr. 2005; Baker et al. 1997]: roughly speaking, around 1 in 1,000 of the population.

1.2 Presenting Web Information in Sign Language

At present there is little access to signed information on the Internet. Relying on text can present great challenges for some deaf people and is a particular problem in the context of eGovernment sites. The information on such sites may be complex. In addition, it is important that people understand it correctly in order that they can successfully access these important services as and when needed. The access that deaf people have to this type of essential information could be greatly improved by the provision of sign language information online.

Sign language can be displayed on Web sites using video, and this method is usually satisfactory to deaf people, especially those who have access to broadband Internet connections. A good example is the British Deaf Association's Web site[1]. However, there are disadvantages to this means of providing information.

- The production cost of video to a professional standard is nontrivial.
- There are continuity issues. Making videos consistent, that is, using the same signer, in the same clothing and with the same background so that signed phrases can be joined together complicates the content maintenance process.

[1]See `http://www.signcommunity.org.uk/`.

- Each time any detail of the content changes, new videos must be made, increasing the costs.
- Storing and downloading videos can also be problematic as they are large files. For home users on dial-up connections, the time and cost involved in downloading video sequences may be prohibitive. This issue will become less significant with increasing penetration of broadband and improvements in video compression technology.
- A sophisticated server architecture is required especially for on-the-fly generation of streamed content.

Instead, we use a virtual avatar driven by animation software, which generates motion data in real time from a scripting notation designed for describing signing. This method does not have the shortcomings just mentioned, and thus provides an attractive alternative to video as a means of presenting signed information on the Internet.

- Both the definitions of signs in our scripting notation and the composition of sequences of signing can be done by one person on a desktop computer. No video or motion capture equipment is required.
- Continuity is not a problem; any piece of signing can be animated for any avatar.
- Details of the content can be edited without having to rerecord whole sequences.
- Bandwidth and disk space demands are negligible. Software is installed on the client machine for translating scripting notation to motion data and rendering the avatar. The scripting notation is all that needs to be transmitted and can be very much smaller than even highly compressed video or motion data. Although we currently use an XML-based scripting notation which tends to verbosity, a binary encoding could easily be created which would require transmitting no more than tens of bytes per second.

Virtual signing has some further advantages.

- The data transmitted is independent of the client-side frame rate which can be set to whatever speed the client machine is capable of rendering the avatar. A user of the system who upgrades their machine with faster graphics would immediately be able to render signing with more frames per second without any changes necessary at the server.
- The user has extra control that is not possible with video. The view angle can be continuously adjusted during playback. The user can switch avatars at will because the signing notation is independent of the avatar.

It must be stressed that the eSign technology presented here is not intended to replace human interpreters. Human skills are still needed to translate spoken and written language to sign language, and the vitality of virtual signing does not currently approach the expressiveness of human signing. However, this approach can make it practical to provide information in signing that otherwise would not be made available.

1.3 Development of Synthetic Animation of Sign Language

The work reported here arises from the international project eSign[2], funded by the EU Framework V programme, which is a successor to some previous projects which we shall briefly describe.

A project known as Simon the Signer [Pezeshkpour et al. 1999] was commissioned by the Independent Television Commission, former regulator of commercial television channels in the UK, to develop technology that would take closed-captioned subtitles from the teletext stream and deliver virtual human signing using Sign-Supported English. The technology involved first parsing and summarizing the text stream since the rate of signing is too fast for many users if the full text is signed. Then motion-captured animations for the individual signs corresponding to the chosen words were presented in a smoothly animated sequence that would blend seamlessly between signs using interpolation techniques. Although the system was technically successful, users expressed a strong preference for the use of true British Sign Language.

A later system, TESSA, developed for the UK Post Office, provided a counter clerk with a tool for communicating with deaf customers. A tuned speech recognition system identified phrases used in communication with Post Office customers and presented the equivalent sign language sequence on a screen facing the customer [Cox et al. 2002, 2003]. Phrases were developed using motion capture and were presented in British Sign Language. In addition to static phrases, blending techniques were used to generate adaptable phrases, including appropriate days of the week, numbers, and prices, depending on context.

The EU project, ViSiCAST[3], explored further use of virtual signing for broadcast applications and for face-to-face interaction as well as on the Web. For most of the applications, motion capture techniques were used to record the body, hand, and face gestures of a deaf signer providing low-level animation data that could be replayed on a client machine. The motion capture process is time-consuming and depends on the use of expensive equipment operated by expert personnel. A significant amount of postprocessing is required, but the result is highly realistic animation that retains much of the natural quality of the original signing. It is possible to replay content captured on different occasions using the same virtual human character, and it is possible to blend between signing sequences so that new content can be created from a database of signs. Additional signs can be added but cannot be created by the average user. It is difficult to make even small changes to signs such as to the facial expressions that often modify the meaning. Different versions of signs are needed for each virtual human character that is used.

For broadcast applications the bandwidth of low-level animation data is relatively high but considerably less than for video, providing the opportunity for closed captioning on a range of channels without consuming unacceptable channel capacity. Further work with the BBC is making progress towards a

[2]eSign: Essential Sign Language Information on Government Networks. `http://www.sign-lang.uni-hamburg.de/eSIGN`.

[3]ViSiCAST: Virtual Signing: Capture, Animation, Storage, and Transmission. `http://www.visicast.co.uk`.

practical system for broadcast signing. ViSiCAST made some early progress on synthesizing character motion from gesture notation which provides a very much more compact representation.

The main focus of this article is the process developed for signed content creation by the eSign project which focussed on Web delivery of eGovernment information. eSign exploits synthesized signing rather than motion capture. An XML-based scripting notation, Signing Gesture Markup Language (SiGML), is used to describe the signing to be performed, along with additional information such as speed and viewpoint. A low-level animation data stream is generated at runtime and used to control a 3D virtual human character [Elliott et al. 2004, 2005; Zwitserlood 2005]. A number of Web sites were developed, animating sign language for each of the partner countries: Deutsche Gebärdensprache (DGS), Nederlandse Gebarentaal (NGT), and British Sign Language (BSL). We report on the tools and content creation approach developed in eSign.

1.4 Related Work

A number of other projects generate animation of sign language using virtual humans. The best-known system for American Sign Language is SigningAvatar, a commercial product from VCom3D[4], for performing sign language on Web pages. It has some similarities with the eSign approach. Synthesis is based on sign representation using the Liddell/Johnson system encoded in XML, but can also include motion capture data or keyframed animation especially in the case of complex signs which cannot be described easily in the Liddell/Johnson system. Animation is performed in VRML. A system for automatic translation from English to signing is provided that targets Signed English rather than ASL. VCom3D has also applied their animation technology to other areas such as lessons in foreign languages and cultural familiarization.

SignSynth [Grieve-Smith 2001] is a signing animation system based on the Stokoe notation and also addressing the issue of translating English text to sign language. The GESSYCA system [Lebourque and Gibet 1999a, 1999b; Gibet and Lebourque 2001] is concerned with French Sign Language. It uses a notation system developed for a project called Qualgest (Qualitative Communication Gestures Specification). THETOS [Francik and Fabian 2002; Suszczanska et al. 2002][5] is a system for animating Polish Sign Language and of automatically translating Polish into PSL. The Auslan Tuition System [Wong et al. 2003; Yeats et al. 2003] uses a signing avatar for teaching signing. A number of systems implement automated animation of fingerspelling where written words are presented using signs for each of the letters [Adamo-Villani and Beni 2004].

The next section provides an overview of the content creation process and delivery over the Internet. We then discuss some details of the content creation tools and the techniques involved in synthesis of signing. We report some results on the usability of a number of Web sites that contain animated signed content.

[4]`http://www.vcom3d.com`.

[5]`http://sun.iinf.polsl.gliwice.pl/sign/`.

2. PRODUCTION OF SIGNED CONTENT

The process of creating and delivering signed content on Web pages involves a number of stages. If the content is originally formulated as written text, it is necessary to determine the appropriate way to present the information through signing. The chosen signing can be represented by a sequence of glosses, which are spoken language words used as labels for the corresponding signs (sometimes parameterized with information that may vary from one use of a sign to another). Each gloss is associated with the SiGML notation that can be used to synthesize the necessary animation. The SiGML may be embedded in a Web page and used as input to a plug-in that performs animation.

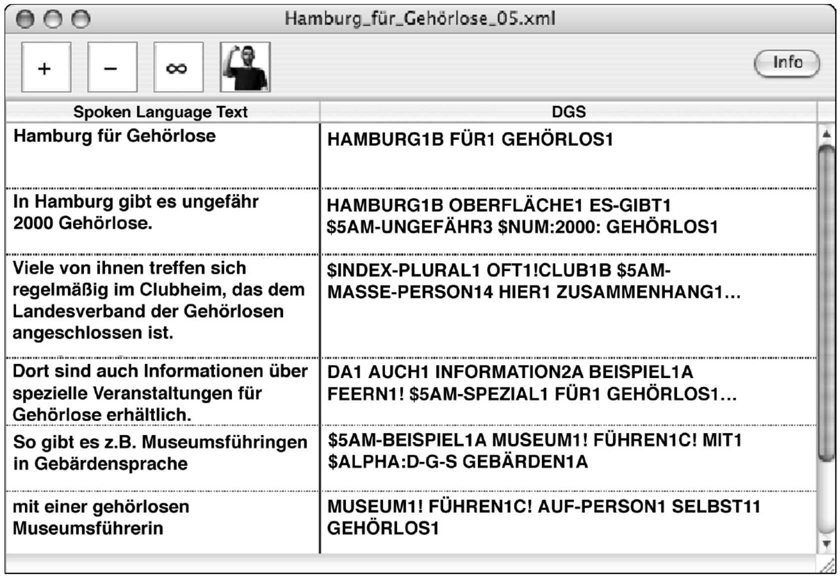
2.1 Choice of Signed Content

Sign languages such as ASL (American Sign Language) and BSL (British Sign Language) are as different from spoken languages as spoken languages are from each other. There exist direct encodings of spoken language, such as Signed English (SE) in the US and Sign-Supported English (SSE) in the UK, which reflect the grammar of the spoken language but use signs from the true sign language. Although converting English text to systems such as SE or SSE is relatively straightforward, these require understanding of the grammar of spoken languages and have not found favor in deaf communities as was reported for Simon the Signer [Pezeshkpour et al. 1999].

Providing signed versions of existing content in a manner acceptable to the intended users therefore requires translation, for example, from English into ASL or BSL. This is as difficult a task as translating from English into French and needs to be undertaken by an expert in the target sign language. (In fact, the task is more like translating from written English into spoken French as issues of prosody must also be addressed.) When complex content is to be signed, such as for eGovernment sites, it may be necessary to adapt and enhance the information to present it in a framework suitable for signing, especially when technical terms do not have a commonly understood equivalent in sign language.

Some progress on automatic translation from English to BSL for a highly constrained domain of discourse was made in the ViSiCAST project [Elliott et al. 2000; Sáfár and Marshall 2001b, 2001a]. However, the translation process cannot be expected to be fully automated in the near future, and it is no more realistic than a proposal that the development of French Web pages should be done by writing in English and then using automatic translation. While translation services such as Babelfish are useful for obtaining the gist of a Web page written in an unfamiliar language, their quality is not yet good enough to serve as a primary means of access.

For sites containing free text that changes frequently, such as news or chat sites, translation to sign language is a bottleneck whatever the technology. Translation to signing systems such as SE or SSE would allow automatic production of new signing content but these systems do not fully address the wishes of deaf people. For those whose first language is BSL, Signed English may be no easier to read than written English. The eSign project and its predecessors



Spoken Language Text	DGS
Hamburg für Gehörlose	HAMBURG1B FÜR1 GEHÖRLOS1
In Hamburg gibt es ungefähr 2000 Gehörlose.	HAMBURG1B OBERFLÄCHE1 ES-GIBT1 $5AM-UNGEFÄHR3 $NUM:2000: GEHÖRLOS1
Viele von ihnen treffen sich regelmäßig im Clubheim, das dem Landesverband der Gehörlosen angeschlossen ist.	$INDEX-PLURAL1 OFT1!CLUB1B $5AM-MASSE-PERSON14 HIER1 ZUSAMMENHANG1...
Dort sind auch Informationen über spezielle Veranstaltungen für Gehörlose erhältlich.	DA1 AUCH1 INFORMATION2A BEISPIEL1A FEERN1! $5AM-SPEZIAL1 FÜR1 GEHÖRLOS1...
So gibt es z.B. Museumsführingen in Gebärdensprache	$5AM-BEISPIEL1A MUSEUM1! FÜHREN1C! MIT1 $ALPHA:D-G-S GEBÄRDEN1A
mit einer gehörlosen Museumsführerin	MUSEUM1! FÜHREN1C! AUF-PERSON1 SELBST11 GEHÖRLOS1

Translation:
—Hamburg for the deaf.
—In Hamburg, there are about 2000 deaf people.
—Many of them meet regularly at the Deaf club, which is affiliated with the Hamburg Association of the Deaf.
—Information is available there about special events for deaf people.
—For example, there are museum tours in sign language
—with a deaf museum guide.

Fig. 1. Text and gloss sequences for some DGS signing.

have been concerned only with sign languages, and not sign-encoded spoken languages.

2.2 Assembling Signed Content

Once an author has determined the signing to be used, it can be represented as a sequence of glosses or words that stand for signs and are written in sign language order. The eSign Editor developed by the IDGS (Institut für Deutsche Gebärdensprache at the University of Hamburg, one of the eSign partners) maintains a lexicon or database of signs, keyed by their glosses. Each gloss is associated with notation for describing a sign. Sequences of glosses can be constructed with the tool illustrated in Figure 1, and the corresponding animation can be viewed for evaluation. Finally, the SiGML script for a sequence can be extracted for use on a Web page. Since SiGML is an XML application language, the script can be embedded in a Web page and manipulated using JavaScript.

For static Web pages with content which is arbitrary but does not frequently change, manual translation as described is sufficient. Minor changes can be incorporated easily without recreating content from scratch as would be necessary with video.

For some applications, it is desirable that Web pages have frequently changing content but confined to a limited grammar and vocabulary, for example, a Web site displaying a daily weather forecast, or one displaying the results of database queries that have a predictable structure. For content of this type, we can know in advance the structure of the source data and the required sign language version. It is feasible to devise a linguistic model of the limited grammatical repertoire and use it to automatically generate the translation into signing of all potential content. An additional benefit of such applications of structured content is that materials can be created by people with no knowledge of sign language: they may select from options presented in written language that will be used to the generate the corresponding sign language based on predetermined transformations. The approach is equally applicable to the generation of high-quality spoken language.

2.3 Delivery of Animation

Signing animation is delivered to the end-user through a Web browser plug-in, which contains an avatar and the software to translate SiGML to motion data. This is packaged as an ActiveX control which the end-user downloads once and installs[6]. Thereafter it will be invoked on signing-enabled Web pages through its JavaScript interface. Scripts, usually attached to buttons on the Web page, are programmed to deliver SiGML to the plug-in, which is then animated using the avatar.

The JavaScript interface supports repeated playing, single stepping, and varying the speed of playback. It returns events to record rendering of animation frames and will deliver the gloss name of the sign played. Standard scripting techniques can be used, for example, to provide controls for the speed of signing. In this way, video player controls and progress bars have been provided on pages using the plug-in. The user is also able to use the mouse to change the viewpoint of the avatar, rotating and tilting the character, moving it up and down, and zooming in or out. The viewpoint can also be controlled by the SiGML script if desired.

According to the way the Web page has been programmed, the avatar can appear within the body of a page, within a subframe, or in a pop-up window. As the plug-in takes a few seconds to load, steps are taken to ensure that it remains resident in memory while the user views multiple signed Web pages. The plug-in provides a panel in which the virtual signer will appear. The designer of a Web page is free to position the panel in the best way to meet the needs of deaf users. On pages designed mainly for use by deaf people, the virtual signer will always be present. Alternatively, signing can be added to a page also used by hearing people in which case the page must be adjusted to accommodate the signer or a pop-up window must be created. We do not have Enough experience with Web pages containing signing to have established clear user-interface principles for virtual signing but early conclusions are presented in a later section.

The present plug-in is provided as an ActiveX object and is therefore available only for Internet Explorer on the Windows platform. A Java version is under

[6]See: `http://www.sys-consulting.co.uk/downloads/esign/`.

development so that the technology should soon be available on a range of browsers on a range of platforms although it may be necessary to use simpler virtual human characters to achieve good performance.

If the plug-in is not installed, or the browser is not capable of supporting the plug-in, it is possible to arrange to fall back to video content although user control of the presentation will be lost and there may be significant delays while videos are downloaded. An example Web site using this approach illustrates an example of the early work on automatic translation from English to BSL[7].

3. TOOLS FOR CONTENT CREATION

The eSign project developed a range of content creation tools. When new content is to be signed, a lexicon database is searched for the individual signs to be used. If the database does not yet contain the sign, it must be created using the gesture notation and entered in the database. Once all signs are available, sequences can be built, tested, and exported for use in Web pages. Both creation of new signs and construction of signed sentences can be performed using the eSign editor. To improve the workflow for creating content that is changed frequently but which follows a predetermined pattern, the Structured Content Creation Tool can be used to build applications that automate much of the work of constructing Web pages.

3.1 Lexicon Development

The SiGML notation is derived from the Hamburg Notation System (HamNoSys), which was developed by the IDGS for recording signs in sign language research [Prillwitz et al. 1989; Hanke 2002]. We describe it in further detail in Section 4. Like most signing notations, it describes signs in terms of hand shapes, hand positions, hand orientations, and movements.

Although HamNoSys also contains provisions for specifying nonmanual movements (an important part of signing, especially facial movements), we found it more suitable to devise a separate and more detailed system for these in SiGML [Elliott et al. 2004]. Nonmanual movements divide into movements of the various parts of the face (mouth, eyes, forehead, etc.), the whole head, the shoulders, and the torso. Mouth movements are particularly important in signing and are of two sorts: *mouthing*, movements that imitate spoken words, and *mouth gestures*, other movements of the mouth such as puffing the cheeks. Mouthing is expressed in SAMPA [Wells 1997], which encodes the phonemes of the word that is mouthed.

HamNoSys is designed to be capable of transcribing the gestures of any sign language in a notation which takes a trained practitioner only a minute or two to record a single sign. It is not a real-time shorthand nor intended as a written form of signing. Its purpose is simply to record the physical postures and movements. The IDGS has a database of some thousands of transcriptions of individual signs. The notation is written with a font of about 250 characters, and an example is shown in Figure 2.

[7] `http://www.visicast.cmp.uea.ac.uk/Demos.`

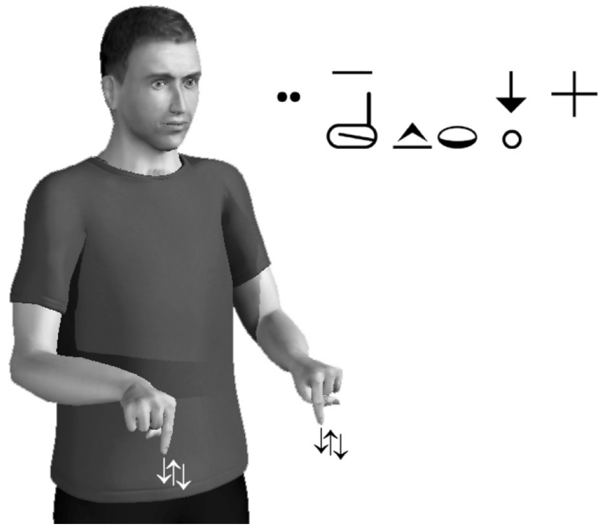


Fig. 2. BSL sign for "here" with HamNoSys transcription.

HamNoSys was originally developed for people to communicate with other people about signing. Its grammar is, as a result, over-elaborate for the needs of computer processing. SiGML expresses the same information in a form more suited to the present application. SiGML also contains provisions for additional information, such as the speed or duration of signing, and the camera position for displaying the avatar. It also allows embedding of low-level motion data. (Its DTD is available at `http://www.visicast.cmp.uea.ac.uk/sigml/sigml.dtd`.)

To enable a new sign to be entered into the lexicon, the eSign editor provides a component for creating HamNoSys descriptions of signs as illustrated in Figure 3. It allows HamNoSys to be either typed on the keyboard or entered via a set of palettes in which the HamNoSys glyphs are laid out in functional groups. Mouthing is entered in SAMPA. The editor can also play a preview animation of the current description. This is necessary because, even for experienced users of HamNoSys, it is very easy to write down an idea of the postures and movements which does not accurately describe their actual geometry. The animation software will play what was written not what was intended. When the HamNoSys description is satisfactory, it is saved in the lexicon database.

3.2 Sign Sequence Construction

The second component of the editor program assembles individual signs into signed sentences and texts. The original spoken language text is split into sentences that will become separate sign sequences. For each sequence, a list of glosses is entered in the correct order to make a grammatical signed sentence, and their HamNoSys definitions are taken from the lexicon database. Prosodic information can be added at this point by adding nonmanual information. There are facial markers for eye gaze, eyebrows, eyelids and nose, and markers for head, shoulders and upper body, various combinations of which have separate grammatical meanings (such as a head shake for a negated sentence, furrowed

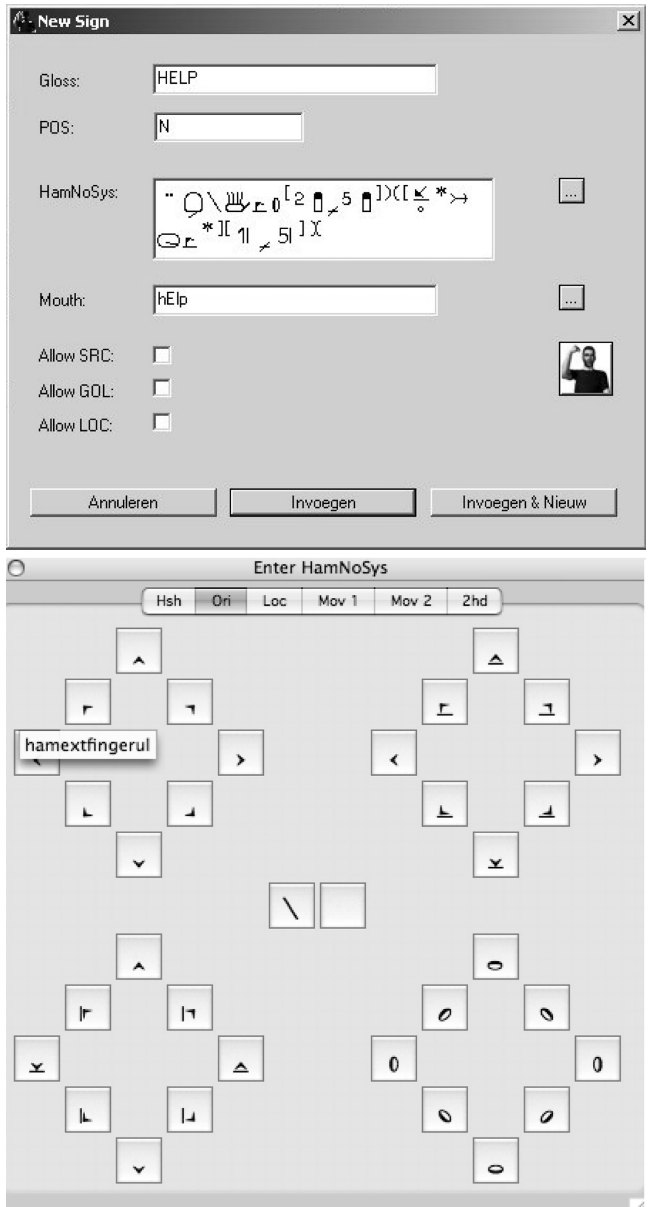


Fig. 3. HamNoSys-based sign editor (Dutch version) showing input palette for hand orientation. The buttons in the upper window are labeled Cancel, Insert, and Insert and New.

eyebrows and head tilted forward for wh-questions). The mouthing which is part of the definition of each individual sign can also be modified at this stage. Figure 4 illustrates the sentence editor.

A number of signs vary according to the context in which they are used. For example, directional verbs will involve a targeted movement where the

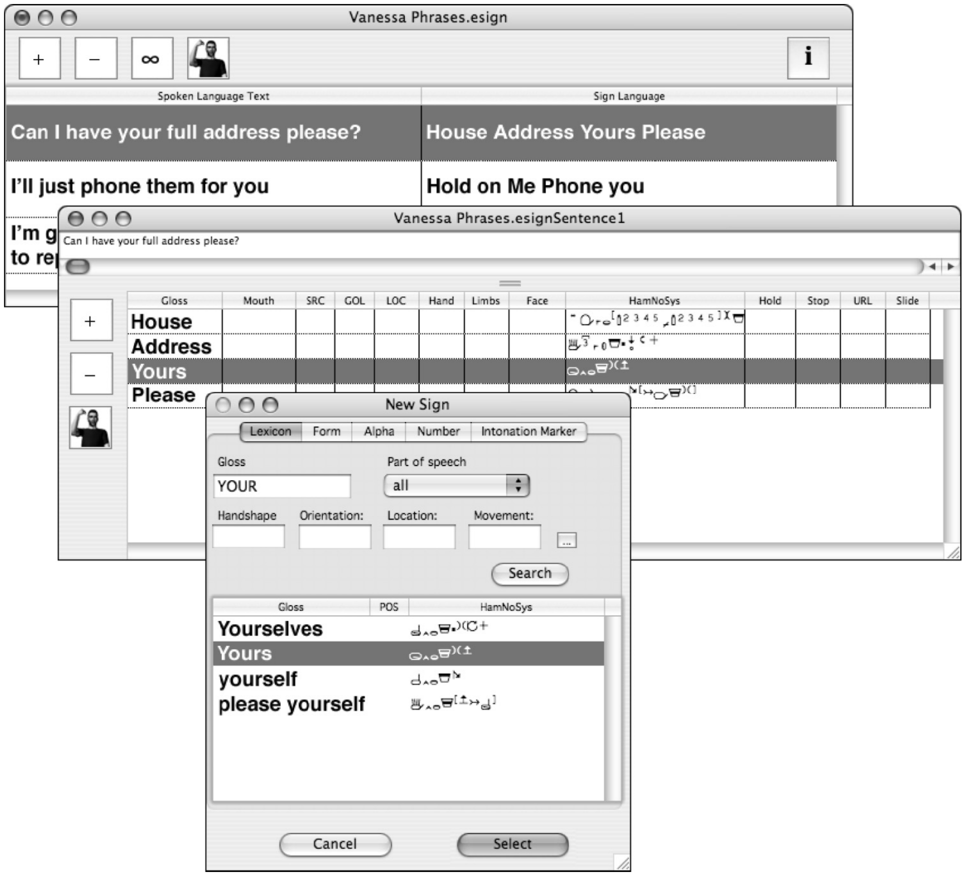


Fig. 4. The sentence composer. The titles of the two upper windows relate to the files being edited.

target depends on the subject or object of the verb. The description of the sign is associated with a default target, but the target may be specified explicitly in which case a part of the HamNoSys description will be modified to reflect the required direction of movement. For some verbs, the hand shape is also varied according to the nature of the object concerned. Signs which require customization according to context are flagged to indicate the attributes that are to be provided. For each attribute that may require modification, an editing dialog or palette can be invoked. A preview of the animation is always available. Figure 5 gives an overview of the facilities available for advanced sequence construction.

It is straightforward for a person skilled in sign language to construct a signed text using the Editor program. Unlike creating content using video, creating animated content is similar to writing a text using a word processor. Content can be previewed using the avatar, mistakes can be repaired immediately, and later adaptations to the signed text can easily be made.

The Editor can be used for the translation of a written text provided that the person creating the text is bilingual. It can also be used for the creation

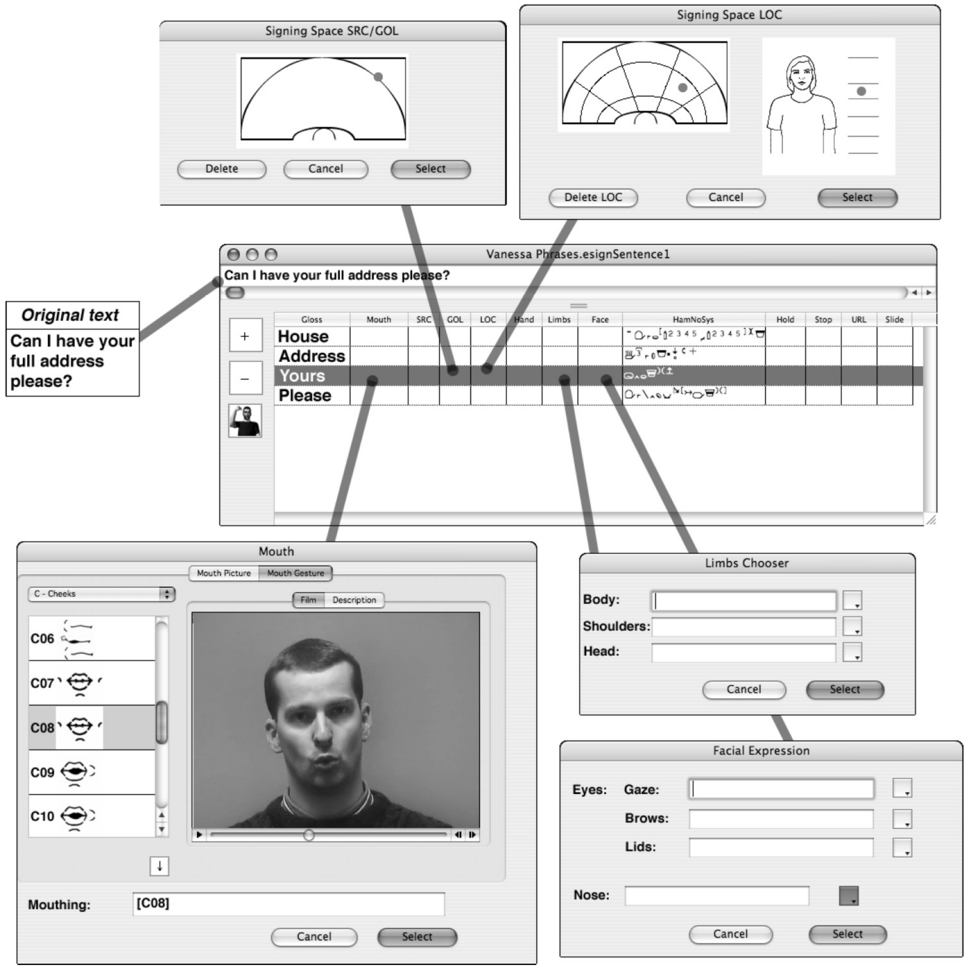


Fig. 5. Sentence editor overview.

of a signed text from scratch by a skilled signer. During the eSign project, the deaf staff in the Netherlands were asked to create signed content for presentations of the technology to the target group. In a follow-up project in the Netherlands (Gebarennet or Sign Web), a course was constructed for deaf signers in which they could learn to create their own content from a database of preconstructed standardized NGT signs and to build this into a Web site. The course members were quite enthusiastic and were only hampered in creating Web sites with animated signing by the limited number of signs in the database.

3.3 Semi-Automatic Content Creation

A translation tool was developed in Visual Basic for the ViSiCAST project, addressing the domain of weather forecasts. A second tool was developed during the eSign project allowing signed job vacancies to be created. Tools such as these contain words and preconstructed phrases from the limited domains concerned in both the spoken and the signed language. A person does not need to know any sign language to use the tools but can construct a text in, for example, Dutch by choosing particular words and phrases. The words and phrases in that text are automatically matched with the correct signs and phrases of the target sign language, and the signed text (e.g., in NGT) can be signed by an avatar.

The weather forecast creation tool provided user interfaces in the English, Dutch, and German languages. Using any of the interface languages, pages with signs (and optional text) for each of the three national sign languages could be generated. In the Netherlands, weather reports were provided on the Web site of a national organization for the deaf (Dovenschap). The weather forecast was refreshed daily during an eight-month trial.

For the signed texts created by the weather forecast tool, motion-captured signs and phrases were used, whereas the signs and phrases in the job vacancy tool were fully synthetic. The output of the content creation tool is a Web page in HTML which contains the scripting to control the avatar plug-in and the SiGML data that will be used to animate the particular weather forecast or vacancy page. The SiGML content is constructed using a template that combines signs according to selections made during content creation. All the signs that could be used are prepared in advance using motion capture or the eSign Editor.

The eSign job vacancy tool is based on the Structured Content Generation Tool (SCGT) which is a Java program that provides a series of panes, radio buttons, and pull-down lists that are used to choose the options for a particular vacancy. The creation tool is illustrated in Figure 6. The six sentences in the lower part of the dialog were constructed with the first six tabs, whose labels mean in English: job title, location, department, qualifications, workload, and start date. The seventh tab, currently active in the figure (contractduur: contract duration) lets the user specify the type of position (tijdelijk: temporary, or vast: permanent), and in the former case, the termination date. Pressing the toevoegen (add) button will generate a new sentence from the user's choices and add it to the panel below. The eighth tab specifies the contact information for the position, and the final tab is for generating an HTML page with embedded signing in Dutch Sign Language. The generated page is shown in Figure 7. Clicking on the icons (Klik op de plaatjes. . .) plays the corresponding animation.

The content creation tool is language-independent. The structure of the user interface and the list of options available is specified in an XML file. Different applications can be produced by developing different XML control files. For example, a reimplementation of the ViSiCAST weather forecast creator, originally written in Visual Basic, was produced using the Java SCGT.

4. ANIMATION OF SIGNING

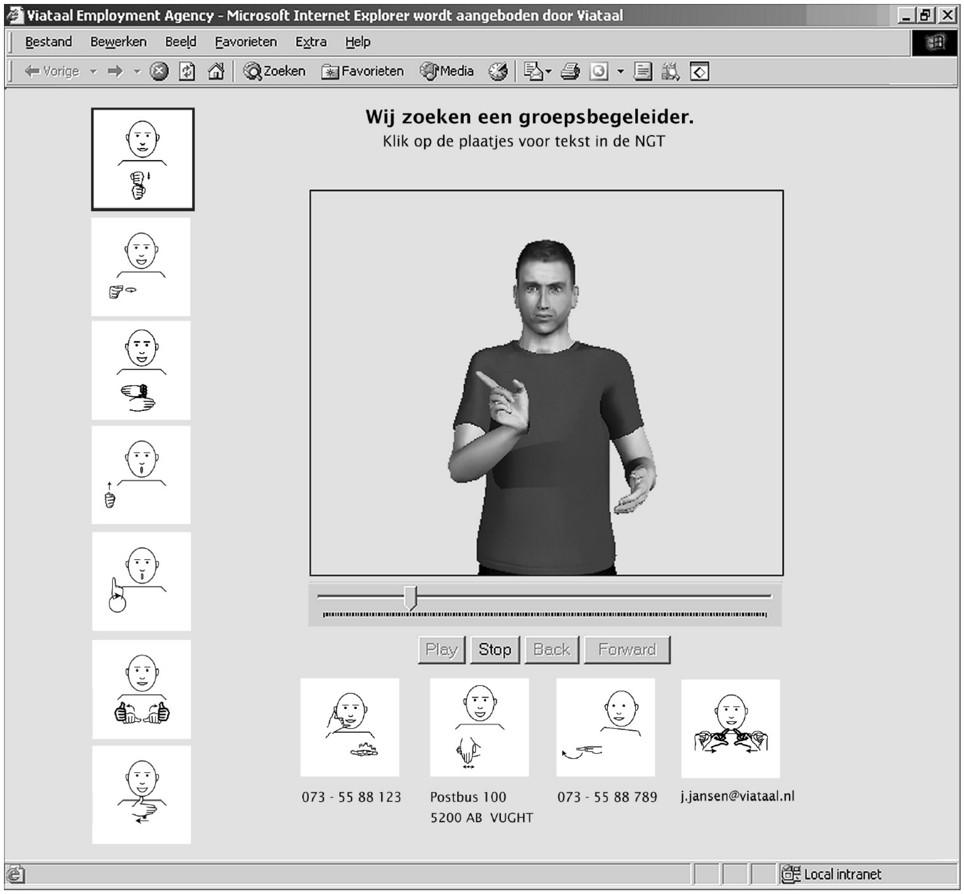
Synthetic animation of signing must start from a description of how each sign is performed and produce frame-by-frame motion data for the avatar.

The sentences mean:
We are looking for a group leader.
The work will be carried out in Nijmegen.
The work will be carried out in the Department of Supervised Autonomous Living.
You must have a diploma, for example the MDGO (a Dutch qualification relating to social work)
The work is for 20 hours per week.
The work begins on 1 September 2004.
See main text for a description of the dialog controls.

Fig. 6. Translation tool for creating signed content (in Dutch).

HamNoSys and SiGML are avatar-independent: they make no reference to the dimensions of a particular avatar. The components of a gesture are described in terms of avatar-relative categories of which we shall illustrate a few examples.

Hand shapes are divided into twelve basic types, illustrated in Figure 8. Each of these can be modified by further notations for individual fingers and the thumb. Hand positions are defined relative to a repertoire of named locations on the surface of the body and in the space immediately in front of the body (*signing space*). The points on the body are illustrated in Figure 9. A location can also be specified as the midpoint of two named locations. It is also possible to specify the spatial relationship of the hands to each other plus the location in space of the pair of hands together (called a *hand constellation*). Hand orientations are described in terms of the direction of the axis of the hand (the direction the fingers would be pointing in if they were straight) and the direction in which the palm faces. Directions are up, down, left, right, inward, outward, and combinations of these. Movements through space can be described as straight, curved, circular, or to a target location; a movement can also be a change of hand orientation or hand shape. Sequential and parallel combinations of movements



The text at the top reads:

"We are looking for a group leader.
Click on the icons for text in Dutch Sign Language."

Fig. 7. Job vacancy Web page.

and repetitions of movements can also be described. Movements can be further specified as small, normal size, or large or slow, normal speed, or fast. Curved movements can be shallow, normal, or deep.

Positions are also discretized: in signing space a finite number of heights, distances from the body, and positions from left-to-right are distinguished. A finite number of points on the surface of the body are also named, the proximity to which can be specified as contacting or near.

HamNoSys only records those parts of the posture and movement which are significant to correct formation of the sign. Most signs involve just the hands and the face: the arms do whatever is necessary to place the hands in the desired position and are not notated except for a few signs where the arm position or movement is itself a significant part of the sign.

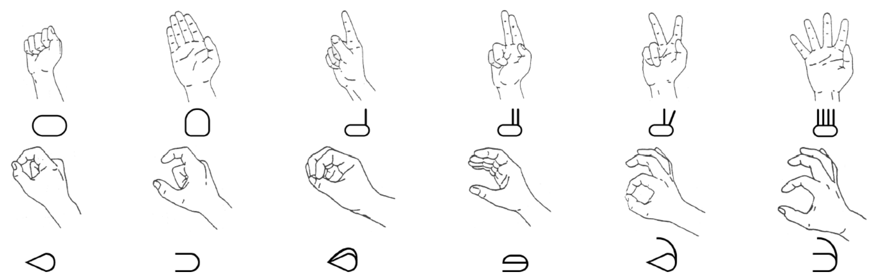


Fig. 8. The basic hand shapes and their HamNoSys symbols.

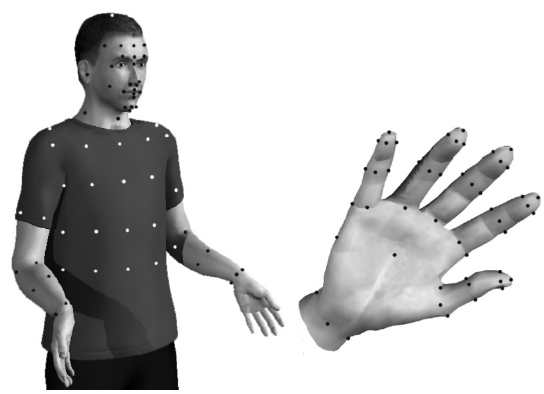


Fig. 9. Locations definable in HamNoSys on the surface of the body.

This discretization and omission of nonsignificant detail allows a typical sign to be notated with between five and twenty HamNoSys symbols, which can be read back and performed faithfully by someone trained in the notation.

4.1 Translation of SiGML to Motion Data

To translate HamNoSys or its XML representation as SiGML into motion data for a particular avatar requires solving the following problems.

(1) The discrete human-meaningful categories such as large and small, fast and slow, locations of points with HamNoSys names, etc., must be translated to numerical quantities appropriate to the avatar.
(2) Using inverse kinematic calculations or otherwise appropriate arm joint rotations and torso movements must be determined to place the hands where required.
(3) All other information that HamNoSys does not record must be inferred. Signs notated for human readers often omit information which is obvious to the human reader but difficult to calculate automatically. For example, contacts between the hands are often notated by the HamNoSys symbol for contact without specifying the particular points on the hands that are to be placed in contact.
(4) Collisions and penetrations of body parts must be avoided.

(5) Facial movement descriptions must be translated to time-varying deformations of the avatar's surface mesh.
(6) All of the postures and movements must be precise enough to be legible as signing, and natural enough to be acceptable to the end-users.
(7) The entire computation must be performed in real time, using a sufficiently small proportion of the time budget to allow for avatar rendering.

4.1.1 *Categories.* How large is a large movement? How curved is a shallow curved arc? How close is near to the chest? What fraction of a second does a fast, slow, or normal speed movement take?

The numbers that we choose for distances are all expressed in terms of dimensions of the avatar. For example, near-the-chest is defined as a certain proportion of arm length since the extent of signing space is determined by the need for the hands to reach all parts of it. Near-the-hand is defined as a proportion of a measurement of hand size. Similar decisions are made for the sizes of movements, the curvature of curved arcs, and all other such categories based on the practice of signing and observation of signers.

The duration of movements in signing is largely independent of the size of the movement but does vary with shape: curved movements take longer than straight ones and circular movements longer still. Each repetition of a repeated movement is somewhat faster than a single occurrence in isolation. Durations were first estimated by counting frames of signing videos, and then adjusted as necessary to produce realistic animation of the right tempo.

The numerical positions of named locations such as point-of-the-chin must be provided to the software as part of the definition of the avatar. Although some of the named points might be computed automatically from the avatar's bones and surface mesh (e.g., the fingertips, and the points on the sides of the fingers), there are many points for which this is not practical. It is up to the person creating the avatar to place these points appropriately and export their numerical positions. This information is an essential part of the avatar description for any application of synthetic animation.

4.1.2 *Arm Joints.* There is a great deal of literature on inverse kinematics algorithms for the human arm [Zhao and Badler 1994; Tolani and Badler 1996; Tolani et al. 2000] and various software libraries such as IKAN are available [University of Pennsylvania 2005]. Given the real-time constraints of the application and taking advantage of the fact that extreme movements are not required for signing, we did not use any of these but adopted a very simple method that gave acceptable results.

The position and orientation of the hand determine the location of the wrist joint. Given the location of the shoulder joint, this determines the elbow angle. There remains one degree-of-freedom of the arm, generally known as the *swivel* angle, a rotation about the line from shoulder joint to wrist joint, which can be specified by the angle between the plane that includes the upper and lower forearm bones and a vertical plane through the shoulder, and wrist joints. When the right hand is in the space directly to the right of the body, we take this angle to be zero, linearly increasing to a value specified for each particular avatar when the hand is directly in front of the shoulder and increasing further as the hand moves to the left.

When the reach of the arm extends more than a certain proportion of the arm length, the sternoclavicular joint is also allowed to rotate to move the shoulder closer to the target. For longer reaches, body rotation and tilt could be added, but this has not yet been implemented.

4.1.3 *Inference of Implicit Information.* In many HamNoSys transcriptions, the hands are specified as being close to each other. However, there is no explicit information in the transcription about the direction from one hand to the other, and there is no way to make this information explicit in HamNoSys. Human readers can easily determine the intended relationship, but it is a nontrivial task to automatically decide whether the hands are to be side-by-side, one above the other, or in some other relationship. In SiGML we have added an attribute to the relevant XML element to specify this information, but this does not solve the problem for HamNoSys. We have implemented various rules to make appropriate choices in different situations, but this is an inherent problem which does not admit of a solution short of extending HamNoSys.

A similar problem occurs with changes of hand orientation: HamNoSys does not have the ability to specify the center of rotation. In some signs, the hand turns from the wrist, in others it is the center of the hand that remains fixed. The best that one can do within HamNoSys is to devise heuristics for determining the appropriate center of rotation.

Some transcriptions specify that a hand is at a particular location without specifying which point on the hand is at that location. In this case, the missing information can be expressed in a more detailed HamNoSys transcription. An example would be a sign in which a hand points inwards and is in contact with the chest. To the human reader, it is obviously the tip of the index finger which is to be in contact. To a program, it is difficult to determine this without implementing computationally expensive collision detection, which, due to the real-time constraint, we do not want to do. To deal with the existing corpus of HamNoSys transcriptions, which often leave out this information, various special-case rules have been implemented. As the last resort, if these rules make the wrong choice, the choice must be made explicit in the HamNoSys transcription.

A significant part of the programming complexity (although not of the execution time) of SiGML implementation is due to the necessity of dealing with such transcriptions which, though sufficiently complete for the human reader, are incomplete from the point of view of software interpretation. In hindsight, the best way to organize this part of the software would be to separate the task of filling in the required information from the task of animation by implementing the former as a transformation from SiGML to SiGML. This would allow the synthesis software to assume that its input was a fully specified SiGML description.

4.1.4 *Collisions and Penetrations.* The software assumes that any SiGML it is given describes a feasible action and, therefore, makes no attempt to avoid certain types of collision. For example, if a hand is notated as being in contact with the chest and then moving inwards, the avatar will simply push its hand through its chest because that is exactly what was asked. In other cases, some effort must be made in the software. For example, one constraint on the arm swivel angle is the requirement that the upper arm not penetrate the chest. This can be ensured by modifying the swivel as necessary, given the location of enough points on the chest to give a sufficiently good approximation to its shape. The points already available in HamNoSys are enough to do this; detailed knowledge of the surface mesh is not required.

4.1.5 *Facial Movements.* The avatar is assumed to be provided with a set of deformations, called *morphs*, of the part of the surface mesh that forms its face (a standard technique of computer animation). The repertoire of facial movements required by SiGML is implemented by defining each one as some mixture of the avatar's morphs together with a temporal profile defining the onset, hold, and release times.

4.1.6 *Naturalism and Precision.* Linear interpolation of movement over time yields highly unnatural-looking movement for living creatures. Natural accelerations and decelerations must be synthesized. We adopted a simple semi-abstract biocontrol model to create suitable interpolations with sufficient parameters to match each of the finite number of different manners that can be expressed in HamNoSys.

These manners are normal, fast, slow, sudden stop and tense. In addition, we found it necessary to distinguish a further manner, lax, which must be used in nonmeaningful movements. Nonmeaningful movements are those which do not serve to communicate any meaning but merely move the hands from the place where they have completed a sign or a part of a sign to the place where the next significant action will begin. There are two situations in which this happens: these are the transition between two signs when the second sign consists of more than just a static posture, and the return stroke of most repeated movements. HamNoSys can also express a form of repetition in which the return stroke is as significant as the forward stroke, and, in this case, the normal manner is used for both. There is a very noticeable visual difference between meaningful and nonmeaningful movements; there is little difference in duration, but the latter performed with less sudden accelerations and decelerations.

Because there are only a finite number of different manners, we can precompute all the desired interpolations once and use lookup tables when generating animation data. Each lookup table defines a function $f : [0, 1] \rightarrow [0, 1]$: when a proportion p of the duration of a movement has elapsed, the movement has achieved a proportion $f(p)$ of its full amount. We use the same set of tables to interpolate all movements; change of hand position, hand orientation, and hand shape, and the onset and release of facial movements.

All of these tables have the property that $f(0) = 0$ and $f(1) = 1$, ensuring that the initial and final positions of every movement are precisely attained. Signing performed by humans often displays coarticulation, not always reaching the target of one movement before flowing into the next. However, we found that imitating this in our animation degraded legibility without improving the visual appearance. To the human viewer, there is a clear difference between fluency and merely bad and sloppy articulation which we do not yet know how to synthesize automatically.

4.1.7 *Real-Time Computation.* The current implementation can generate 15,000 frames of motion data per second on a 1.7GHz Pentium III PC. Each frame consists of 50 to 80 joint rotations, depending on the avatar, plus some number of facial morph amounts. If the avatar plays at 25 frames per second, this represents 25/15,000 or 0.17% of the time budget. It is thus highly practical and could even synthesize data for multiple avatars simultaneously at higher frame rates.

We have outlined the animation generation software that converts SiGML notation to avatar animation data. The technical design of a prototype version of the system and some designs for a system that would overcome some of the ambiguities in the HamNoSys translation are presented elsewhere [Kennaway 2001, 2003].

5. IMPLEMENTATION OF DEMONSTRATION SITES

The eSign project concentrated on government information. A selection was made of forms that deaf people often need to fill out; these were either translated or provided with signed information. Figure 10 shows an example of such a form. The caption at the top of the Web page reads "Application form for an interpreter". The form is for someone to request the services of a signing interpreter, and the form items let the user specify when and where the service is required, the sign language, and the number of people. Clicking on the caption beside each item (location, begin date, end date, etc.) plays a signing animation of the caption. Existing Web sites were adapted to include pages with animated signing: an information site for Hamburg citizens and visitors[8]; a site giving information about certain labor legislation in the Netherlands; and the Norfolk Deaf Connexions Web site[9]. One of the applications concerned job vacancies at one of the partners: these were provided with a signed summary, illustrated in Figure 7.

For all these applications, the layout and structure of the original pages was kept as much as possible. Sometimes it was necessary to make amendments or even to rebuild the Web site in order to put in as many facilities of animated signing as possible.

In a follow-up project in the Netherlands, a bilingual Web page (Dutch/NGT) was created[10]. This Web site was aimed at deaf signers in the Netherlands, explaining the animation technology and its possibilities. The Web site still provides an information platform.

Viataal[11], one of the eSign partners, is currently working on applying synthetic signing to deaf education in various ways: for instruction (e.g. signed

[8] http://gebaerden.hamburg.de.
[9] http://www.deafconnexions.org.
[10] http://www.gebarennet.nl.
[11] http://www.viataal.nl.

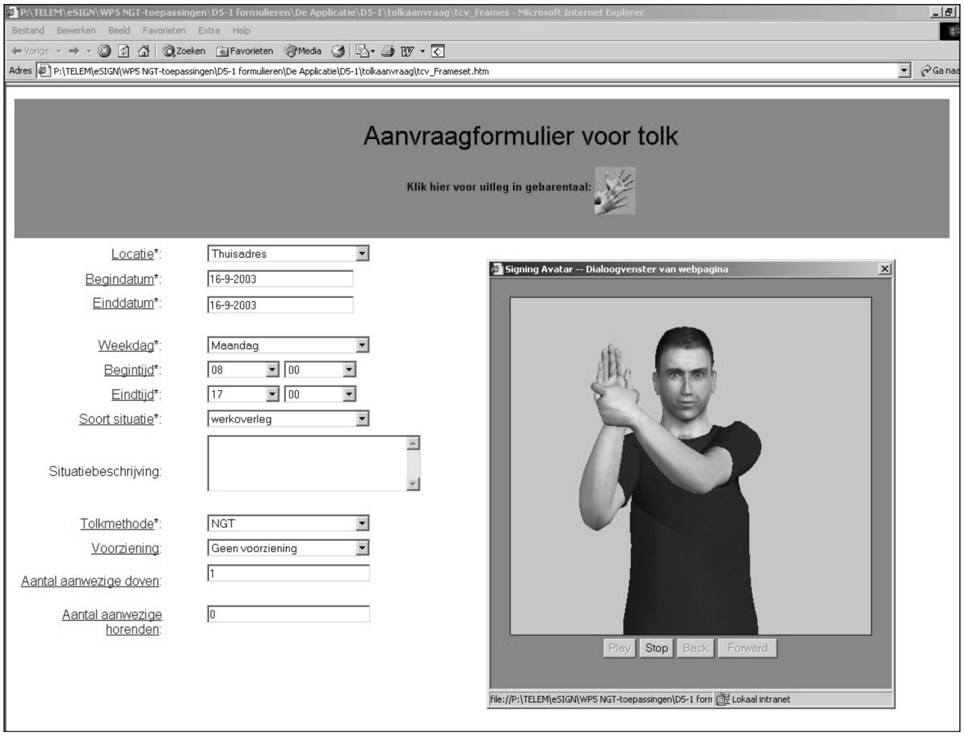


Fig. 10. Avatar providing assistance with a form (on a Dutch Web page).

textbooks), questionnaires, exercises with signed feedback, etc. The use of a signing avatar allows great flexibility in the creation of content.

6. EVALUATIONS

Eight evaluations have been conducted aimed at assessing the level of comprehension of the synthetic signing developed in eSign. The aim was to gain feedback on the acceptability of the avatar, the applications and the design of applications with avatar signing, to use as guidance for further improvement. These evaluations were spread over the course of the project: two evaluations took place in Germany and three in each of the UK and the Netherlands.

Although the tasks set to the participants differed, the evaluations were similar in several respects. All but one evaluation used a lab setting. There were one-to-one sessions with a participant and a signing (usually deaf) evaluator, and the computer configuration necessary for the tasks. In some cases, the sessions were recorded on video. The numbers of participants were in general small due to the size of the target group, ranging from six to ten participants. One evaluation had 15 participants. The ages of the participants varied from 16 to 66. The number of male participants was slightly higher than that of female participants. All participants were good signers, most of them deaf, some hard of hearing, all using signing on a daily basis and as their preferred language.

The evaluations had similar formats, that is, an introduction in which the project aims and the tasks were explained, a short period of familiarization with avatar signing, the particular tasks, and finally a questionnaire/interview. In the sections that follow, the results with respect to sign recognition/comprehension, and user appreciation of (applications with) avatar signing will be described.

6.1 Comprehensibility

Comprehension of the synthetic signing was measured in various ways including whether single signs were recognized and whether isolated sentences and chunks of signed text were comprehended.

The German partner concentrated on comprehension of sentences, whereas in the UK and the Netherlands, comprehensibility of single signs, sentences and short text chunks were scored [Sheard et al. 2004]. Subsequent research in the Netherlands aimed at comparing comprehension of signed movies, avatar signing, and sign pictures [Vink and Schermer 2005], the results of which will be reported here as well.

6.1.1 *Comprehension of Single Signs.* Sign comprehension was tested by asking participants to watch a range of signs presented by vGuido (the avatar appearing in Figure 2) and to repeat each sign after it was played. If necessary, vGuido's sign was replayed. The results of the first evaluation were a comprehensibility score of 70% in the UK and of 75% in the Netherlands. A final evaluation held in the Netherlands, after several improvements with respect to manual and nonmanual components had been made, resulted in a comprehensibility rating of 95%. When asked to indicate whether the signs in 20 presented pairs of signs were the same or different, 90% were correct. These are extremely good scores, and some remarks are appropriate here. Implementation of nonmanual (mainly facial) components presented some difficulties and was in progress until the end of the project. Since it would be unrealistic to score comprehensibility of signs with missing or incorrect components, only recognition of the manual part was scored in all of these tests. In practice, this means that polysemous signs that are only distinguished by mouthing were scored as OK if the participant used a different mouthing than the test sign. For example, the NGT signs for brother and sister have exactly the same manual shape but are only distinguished by different mouthings. When a participant repeated the sign for sister as the sign for brother, this was counted as correct.

Missing or failing to recognize a sign was sometimes due to signs that were unknown to participants because of regional variation (especially in the UK). Other errors can be attributed to deficiencies in the sign transcription or animation quality: the shapes of the hands, movement, or position of the hands, and lack of correct mouthing (UK and NL).

Vink and Schermer [2005] report on an independent sign comprehension task, again rating comprehension by percentage of correct repetitions. In this test, mouthing that obligatorily accompanies signs was taken into consideration (mistakes in nonobligatory mouthing were not counted). Here, 85% of the signs were recognized correctly, indicating that the implementation of mouthing had improved considerably.

6.1.2 *Comprehension of Signed Sentences and Text Chunks.* Sentence comprehension was tested in different ways. The first test, administered in Germany, compared comprehension of synthetic signs (performed by vGuido) and motion captured signs as in ViSiCAST (performed by Visia2, who appears in Figure 6). Participants were shown sentence pairs, one sentence by vGuido, the other by Visia2. They were asked to indicate whether the content of the sentence pairs was the same or not. Comprehension of the synthetic signing was 40.4%. In the UK, recognition of sentences signed by vGuido at that stage was rated at 46% of correctly repeated sentences; this was 35% in the Netherlands. In a test administered in the Netherlands at a later stage, comprehension increased to 58%, and Vink and Schermer [2005] even scored 71% for correctly repeated sentences. The second test administered in Germany in which participants were asked to retell the gist of sentences of one of the applications resulted in 62% of reasonable to good comprehension. Mistakes in comprehension were due to unclear sign shapes (e.g., unclear handshapes or wrong movements), ambiguity, missing or unclear nonmanuals, and missing or incorrect prosody (such as pauses).

In the UK and the Netherlands, further tests measured comprehensibility of longer stretches of signed text. In the UK, participants were shown one of the applications and were subsequently asked a few content questions and then to fill out a form with signed support. The content questions were answered correctly in 71% of the cases, and 5 out of 7 participants filled out the form without any trouble; 2 of them made minor errors. It must be noted here that, besides the signing, English text was available. Participants in the Netherlands were shown an application (without additional text) and given five assignments. All assignments were carried out successfully.

6.1.3 *Summary.* The evaluations show that comprehension of synthetic signing has improved during the project as a result of improvements made to the animation software resulting from the earlier tests and also from informal feedback. However, the project duration was too short to be able to incorporate all of the desired improvements. We are aware of some gaps in the avatar's repertoire of nonmanual bodily behavior and facial animation, and there are some issues concerning manual components that need to be further developed. We believe that with these techniques we would be able to raise the recognition rates for continuous signed text to the levels of above 90% that we have achieved for isolated signs.

6.2 Web Page Design

Different Web page designs were used in order to find out whether signers have particular preferences. The avatar appeared in either a subframe or in a pop-up window in various positions on the screen. Background coloring was varied. Sometimes text was provided alongside the signing, sometimes not. Pages might contain video controls, a scroll bar, or a speed control, similar to those available in most media players. In some pages, the signing was a translation of the written text, in others it was explanatory rather than a translation. Signing could be induced by hyperlinked text (words, sentences or paragraphs) or by particular buttons.

Some general preferences with regard to the design of an application Web page resulted. The participants prefered to be able to see written text as well as signing and were glad to have as much control as possible. Besides that, preferences varied with the particular application. In some cases, a translation was considered fine, whereas in other applications, some participants preferred explanations. There was no general preference for a positioning of the avatar window. Although fixed avatar frame windows were favored over pop-up windows, this was caused by the fact that pop-up windows overlapped the text. There was a slight preference for buttons to invoke signing in the applications presented.

More and different types of applications are needed before general guidelines for signing Web pages can be formulated.

6.3 Subject Appreciation of (Applications with) Synthetic Signing

Most of the participants taking part in the evaluations found it pleasant to see animated signing even though it was sometimes hard to understand. Most participants found the signing somewhat unnatural and complained about the clarity of mouthing and lack of facial expressiveness. Note that in signing, facial expression carries not only emotional color, as it does for speech, but also some of the grammatical structure and semantics.

vGuido has been designed as a realistic-looking male virtual person, and, although a few participants thought his appearance was cool, many indicated that it was too dark, especially in the facial area. It was thought that a shave might improve this. (The issue of darkness was solved to some extent by changing the lighting.) vGuido's face was considered as very serious by some, as unfriendly by others. Some participants found that he looked too broad and muscular ("I don't want to meet him in a dark street at night"). His hands were considered too big and it was indicated that his thumbs were awkward.

Control over the signing in the sense of video buttons, the possibility of resizing and turning the avatar, and controlling the speed of signing was much appreciated. Other features in demand were being able to enlarge the avatar window, change the background color of the avatar window, and having a choice of avatars.

The idea of animated signed assistance was appreciated by many subjects. They thought it could be really helpful, especially for those deaf who have severe trouble reading. Nevertheless, almost all participants appeared not to fully rely on synthetic signing because they indicated that they wanted written text to be present alongside the signing. Besides Web pages, it was mentioned that synthetic signing might be helpful in giving information at train stations, airports, townhalls, hospitals and travel agencies.

6.4 Summary

The evaluations gave good indications of where the synthetic signing technology needed improvement and of the needs and desires of the users of synthetic signing applications. Naturally, the results have been used to drive improvements. Implementation of nonmanual components, especially mouthing, proved to be quite difficult. Nevertheless, the increasing comprehension scores (particularly the sentence comprehension scores) indicate that this has been quite successful, although further improvement is still desirable.

Furthermore, where possible, other aspects (such as lighting, addition of controls) have been adapted in the course of the project. It was not possible to design a wide range of signing avatars within the project; this is something that can be done in the coming years.

There have been numerous proposals for higher-level scripting languages for embodied conversational agents, some of which are surveyed in Pelachaud [2003]. The work reported here could serve as the lower-level means of synthesizing the animations required by such systems. We are also currently exploring this type of low-level animation scripting for other VR applications involving expressive virtual human characters.

7. CONCLUSION

The project has demonstrated the practicality of using computer-generated avatars on Web pages of a standard sufficient to communicate with deaf people through signing. The effort involved in doing so should, when the technology has matured, be no more than that required to produce a translation into another spoken language. By transmitting the signing information in the form of a transcription notation animated on the client, the bandwidth required is far below that of even highly compressed video or of raw motion data and allows user choice of viewpoint and even the virtual character used to present the signing. Beyond the rendering time of the avatar, the overhead of generating motion data from signing notation is negligible. Signing content can be rapidly created without the expense of recording sessions with human signers or any expensive technology.

ACKNOWLEDGMENTS

We acknowledge the support of our colleagues and partners in the eSign project. The vGuido avatar was created by Televirtual Ltd. and animated by software developed under the direction of Ralph Elliott at UEA. The BSL site for Deaf Connexions in the UK was developed by Norfolk County Council, and the Royal National Institute for Deaf People supported by Judy Tryggvason at UEA. Viataal developed several NGT sites using the eSign Editor and the Structured Content Generation Tool from UEA. The DGS site for Hamburg citizens was developed by EDS Business Solutions GmbH and the Institut für Deutsche Gebärdensprache (IDGS), University of Hamburg. The eSign Editor was developed by the IDGS under the direction of Thomas Hanke. Neil Thomas of the RNID provided useful comments on the article. We are grateful for constructive comments and suggestions from referees of this article.

We thank the deaf users of sign language who have provided valuable feedback on the quality of our animations and have followed the progress of the project with interest.

REFERENCES

ADAMO-VILLANI, N. AND BENI, G. 2004. Automated finger spelling by highly realistic 3D animation. *British J. Educat. Tech. 35*, 3, 345–362.

ALLEN, T. E. 1986. Patterns of academic achievement among hearing impaired students: 1974 and 1983. In *Deaf Children in America*, A. N. Schildroth and M. A. Harchmer, Eds. College Hill Press, San Diego, CA, 161–206.

BAKER, A., HENDRIKX, A., KNOORS, H., VAN DER LEM, T., LEVELT, W., SCHADEE, M., AND WESEMANN, J. 1997. Meer dan een Gebaar [More than a sign]. Report on the recognition of Sign Language of the Netherlands (NGT) as an official language. Commissioned by the *Ministry of Education, Culture and Science and the Ministry of Welfare, Health and Sport of the Dutch Government*. The Hague. SDU publishers.

CONRAD, R. 1979. *The Deaf Schoolchild*. Harper and Row, NY, NY.

COX, S. J., LINCOLN, M., NAKISA, M., WELLS, M., TUTT, M., AND ABBOTT, S. 2003. The development and evaluation of a speech to sign translation system to assist transactions. *Int. J. Hum. Comput. Interact. 16*, 2, 141–161.

COX, S. J., LINCOLN, M., TRYGGVASON, J., NAKISA, M., WELLS, M., TUTT, M., AND ABBOTT, S. 2002. TESSA, a system to aid communication with deaf people. In *5th International ACM SIGCAPH Conference on Assistive Technologies (ASSETS)*. Edinburgh, Scotland. ACM, 205–212.

DEUCHAR, M. 1984. *British Sign Language*. Routledge, London, UK.

ELLIOTT, R., GLAUERT, J. R. W., JENNINGS, V., AND KENNAWAY, J. R. 2004. An overview of the SiGML notation and SiGMLSigning software system. In *4th International Conference on Language Resources and Evaluation (LREC 2004)*. Lisbon, Portugal. O. Streiter and C. Vettori, Eds. 98–104.

ELLIOTT, R., GLAUERT, J. R. W., AND KENNAWAY, J. R. 2004. A framework for non-manual gestures in a synthetic signing system. In *2nd Cambridge Workshop on Universal Access and Assistive Technology (CWUAAT)* Cambridge, UK. S. Keates, P. J. Clarkson, P. Langdon, and P. Robinson, Eds. 127–136.

ELLIOTT, R., GLAUERT, J. R. W., AND KENNAWAY, J. R. 2005. Developing techniques to support scripted sign language performance by a virtual human. In *3rd International Conference on UA in HCI*. Las Vegas, NV, Universal Access in HCI: Exploring New Dimensions of Diversity. Vol. 8 (CD–ROM).

ELLIOTT, R., GLAUERT, J. R. W., KENNAWAY, J. R., AND MARSHALL, I. 2000. Development of language processing support for the ViSiCAST project. In *4th International ACM SIGCAPH Conference on Assistive Technologies (ASSETS'00)*. Washington DC. ACM.

FRANCIK, J. AND FABIAN, P. 2002. Animating sign language in the real time. In *20th IASTED International Multi-Conference on Applied Informatics*. 276–281.

GIBET, S. AND LEBOURQUE, T. 2001. High-level specification and animation of communicative gestures. *J. Visual Lang. Comput. 12*, 657–687.

GORDON JR., R. G., ED. 2005. *Ethnologue: Languages of the World*. SIL International, Dallas, TX. http://www.ethnologue.com/.

GRIEVE-SMITH, A. B. 2001. SignSynth: A sign language synthesis application using Web3D and Perl. In *4th International Workshop on Gesture and Sign Language Based Human-Computer Interaction*. Lecture Notes in Artificial Intelligence, vol. 2298. I. Wachsmuth and T. Sowa, Eds. Springer, 134–145.

HANKE, T. 2002. HamNoSys in a sign language generation context. In *Progress in Sign Language Research*. R. Schulmeister and H. Reinitzer, Eds. Signum Verlag, Hamburg, Germany, 249–264.

HOLT, J. A., TRAXLER, C. B., AND ALLEN, T. E. 1996. Interpreting the scores: A user's guide to the 9th edition Stanford Achievement Test for educators of deaf and hard-of-hearing students. Unpublished manuscript, Gallaudet University, Washington DC.

KENNAWAY, J. R. 2001. Synthetic animation of deaf signing gestures. In *4th International Workshop on Gesture and Sign Language Based Human-Computer Interaction*. Lecture Notes in Artificial Intelligence, vol. 2298. I. Wachsmuth and T. Sowa, Eds. Springer, 146–157.

KENNAWAY, J. R. 2003. Experience with and requirements for a gesture description language for synthetic animation. In *5th International Workshop on Gesture and Sign Language Based Human-Computer Interaction*. Lecture Notes in Artificial Intelligence, vol. 2915. A. Camurri and G. Volpe, Eds. Springer, 300–311.

LEBOURQUE, T. AND GIBET, S. 1999a. A complete system for the specification and the generation of sign language gestures. In *3rd International Workshop on Gesture-Based Communication in Human-Computer Interaction*. Lecture Notes in Artificial Intelligence, vol. 1739. A. Braffort, R. Gherbu, S. Gibet, J. Richardson, and D. Teil, Eds. Springer, 227–238.

LEBOURQUE, T. AND GIBET, S. 1999b. High level specification and control of communication gestures: the GESSYCA system. *Comput. Animat. '99*. 24–35.

PELACHAUD, C. 2003. Overview of representation languages for ECAs. Tech. rep., Paris VIII, IUT Montreuil.

PEZESHKPOUR, F., MARSHALL, I., ELLIOTT, R., AND BANGHAM, J. 1999. Development of a legible deaf signing virtual human. In *Proceedings IEEE International Conference on Multimedia Computing and Systems*. Florence, Italy, IEEE, 333–338.

PRILLWITZ, S., LEVEN, R., ZIENERT, H., HANKE, T., HENNING, J., AND COLLEAGUES. 1989. *HamNoSys Version 2.0: Hamburg Notation System for Sign Languages—An Introductory Guide*. International Studies on Sign Language and the Communication of the Deaf, Volume 5. University of Hamburg. `http://www.sign-lang.uni-hamburg.de/Projekte/HamNoSys/HNS4.0/HNS4.0eng/Contents.html`.

SÁFÁR, E. AND MARSHALL, I. 2001a. The architecture of an English-text-to-sign-languages translation system. In *Recent Advances in Natural Language Processing (RANLP)*. Tzigov Chark, Bulgaria, G. Angelova, Ed. 223–228.

SÁFÁR, E. AND MARSHALL, I. 2001b. Translation of English text to a DRS-based, sign language oriented semantic representation. In *Conference sur le Traitment Automatique des Langues Naturelles (TALN'01)*. Vol. 2. Tours, France, 297–306.

SHEARD, M., VAN DER SCHOOT, S., ZWITSERLOOD, I., VERLINDEN, M., AND WEBER, I. 2004. Evaluation Reports 1 and 2 of the EU project *Essential Sign Language Information on Government Networks*.

SUSZCZANSKA, N., SZMAL, P., AND FRANCIK, J. 2002. Translating Polish texts into sign language in the TGT system. In *20th IASTED International Multi-Conference on Applied Informatics*. 282–287.

TOLANI, D. AND BADLER, N. I. 1996. Real-time inverse kinematics of the human arm. *Presence 5*, 4, 393–401.

TOLANI, D., GOSWAMI, A., AND BADLER, N. I. 2000. Real-time inverse kinematics techniques for anthropomorphic limbs. *Graphical Models 62*, 5, 353–388.

UNIVERSITY OF PENNSYLVANIA. 2005. IKAN: Inverse Kinematics using ANalytical methods. `http://hms.upenn.edu/software/ik/ik.html`.

VINK, M. AND SCHERMER, T. 2005. Verslag onderzoek naar het gebruik van de avatar in vergelijking met tekeningen en films van gebaren (Report of research on the use of an avatar compared with drawings or films of gestures). Tech. rep., Nederlands Gebarencentrum (Dutch Sign Language Center).

WAUTERS, L. N. 2005. Reading comprehension in deaf children: The impact of the mode of acquisition of word meanings. Ph.D. thesis, Radboud University Nijmegen.

WELLS, J. C. 1997. SAMPA computer readable phonetic alphabet. In *Handbook of Standards and Resources for Spoken Language Systems*, Part IV, Section B, D. Gibbon, R. Moore, and R. Winski, Eds. Mouton de Gruyter, Berlin, Germany. `http://www.phon.ucl.ac.uk/home/sampa`.

WONG, J., HOLDEN, E., LOWE, N., AND OWENS, R. 2003. Real-time facial expressions in Auslan Tuition System. In *5th IASTED International Conference on Computer Graphics and Imaging*. `http://auslantuition.csse.uwa.edu.au/help.html`.

YEATS, S., HOLDEN, E., AND OWENS, R. 2003. An animated Auslan tuition system. *Int. J. Machine Graphics Vision 12*, 2, 203–214.

ZHAO, J. AND BADLER, N. I. 1994. Inverse kinematics positioning using nonlinear programming for highly articulated figures. *ACM Trans. Graph. 13*, 4, 313–336.

ZWITSERLOOD, I. 2005. Synthetic signing. In *The World of Content Creation, Management, and Delivery*. (IBC'05). 352–357.

Received October 2005; revised April 2006, June/July 2006; accepted October 2006 by Andrew Sears, Vicki Hanson, and Brod Myers