The Workshop Programme

Sunday, May 30th 2004

	Invited talks
9:00 – 9:30	Thomas Hanke *HamNoSys – Representing Sign Language Data in Language Resources and Language Processing Contexts*
9:30 – 10:00	Carol Neidle, Robert G. Lee *Corpus Annotation, SignStream*
10:00 – 10:30	Richard Gleaves, Valerie Sutton *SignWriter*
	10:30 – 11:00 Break
11:00 – 12:30	**Oral presentations (I session)**
11:00 – 11:30	Galini Sapountzaki, Eleni Efthimiou, Costas Karpouzis, Vassilis Kourbetis *Open-ended Resources in Greek Sign Language: Development of an e-Learning Platform*
11:30 – 12:00	Onno Crasborn, Els van der Kooij, Daan Broeder, Hennie Brugman *Sharing sign language corpora online: proposals for transcription and metadata categories*
12:00 – 12:30	Matt Huenerfauth *Spatial Representation of Classifier Predicates for Machine Translation into American Sign Language*
	12:30 – 14:00 Lunch
	Oral presentations (II session)
14:00 – 14:30	Antônio Carlos da Rocha Costa, Graçaliz Pereira Dimuro, Juliano Baldez de Freitas *A Sign Matching Technique to Support Searches in Sign Language Texts*
14:30 – 15:00	Angel Herrero *A Practical Writing System for Sign Languages*
15:00 – 15:30	Maria Papadogiorgaki, Nikos Grammalidis, Nikos Sarris, Michael G. Strintzis *Synthesis of Virtual Reality Animations from SWML using MPEG-4 Body Animation Parameters*
	15:30 – 16:00 Coffee Break

16:00 – 18:00 **Poster session**

Eleni Efthimiou, Anna Vacalopoulou, Stavroula Evita Fotinea, Gregory Steinhauer
Multipurpose Design and Creation of GSL Dictionaries

Chiara Vettori, Oliver Streiter, Judith Knapp
From Computer Assisted Language Learning (CALL) to Sign Language Processing: the Design of E-LIS, an Electronic Bilingual Dictionary of Italian Sign Language and Italian

Rubén Nogueira, Jose M. Martínez
19th Century Signs in the Online Spanish Sign Language Library: the Historical Dictionary Project

Elana Ochse
A language via two others: learning English through LIS

Ingvild Roald
Making Dictionaries of Technical Signs: from Paper and Glue through SW-DOS to SignBank

Steven Aerts, Bart Braem, Katrien Van Mulders, Kristof De Weerdt
Searching SignWriting Signs

Inge Zwitserlood, Doeko Hekstra
Sign Printing System – SignPS

Boris Lenseigne, Frédérick Gianni, Patrice Dalle
A New Gesture Representation for Sign Language Analysis

Jose L. Hernandez – Rebollar
Phonetic Model for Automatic Recognition of Hand Gestures

Daniel Thomas Ulrich Noelpp
Development of a New „SignWriter" Program

Ralph Elliott, John Glauert, Vince Jennings, Richard Kennaway
An Overview of the SiGML Notation and SiGMLSigning Software System

Jan Bungeroth, Hermann Ney
Statistical Sign Language Translation

Guylhem Aznar, Patrice Dalle
Computer Support for SignWriting Written Form of Sign Language

Yiqiang Chen, Wen Gao, Changshui Yang, Dalong Jiang, Cunbao Ge
Chinese Sign Language Synthesis and Its Applications

Paola Laterza, Claudio Baj
Progetto e-LIS@

Workshop Organisers

- Oliver Streiter, *Research Area* "*Language and Law" - European Academy of Bolzano, Italy*
- Antônio Carlos da Rocha Costa, *Escola de Informática - Universidade Católica de Pelotas, Brazil*
- Christian Retoré, *Laboratoire Bordelais de Recherche en Informatique, France*
- Chiara Vettori, *Research Area* "*Language and Law" - European Academy of Bolzano, Italy*

Workshop Programme Committee

- Antônio Carlos da Rocha Costa, *Escola de Informática - Universidade Católica de Pelotas, Brazil*
- Carol Neidle, *Department of Modern Foreign Languages and Literatures - Boston University, Boston MA*
- Chiara Vettori, *Research Area* "*Language and Law" - European Academy of Bolzano, Italy*
- Christian Retoré, *Laboratoire Bordelais de Recherche en Informatique, France*
- Eva Safar, *School of Computing Sciences - University of East Anglia, Norwich, England*
- Ian Marshall, *School of Computing Sciences, University of East Anglia, Norwich, Englan*
- Marco Consolati, *Cooperativa Alba, Torino, Italy*
- Oliver Streiter, *Research Area* "*Language and Law" - European Academy of Bolzano, Italy*
- Patrice Dalle, *Équipe "Traitement et Compréhension d'Images", IRIT - Université Paul Sabatier, France*

Table of Contents

Author Index

Preface

On behalf of the program committee for the LREC 2004 "Workshop on the Processing of Sign Languages", we are pleased to present you with the proceedings which contain the papers accepted for presentation at the Lisbon meeting on May 30th, 2004.

This volume, full of eye-catching signs, symbols, robots and screen-shots may charmingly attract readers who, although having a sound knowledge of Natural Language Processing, might be confused by the great variety of topics and approaches. How do SignWriting, avatars, XML, videos and image recognition fit together? Are they competitive approaches or different solutions to different problems? Where will future research lead us, which endeavours answer real social needs and which scenarios are still illusionary - or congenially visionary?

As always, the answers to these questions lie between slow and quick, up and down, straight and curbed. It is by drawing analogies to the processing of spoken languages that we might better understand the contribution and benefits of the different approaches, span the space of possible research and identify future tendencies in the research on the processing of sign languages.

Trivially speaking, spoken languages are spoken and heard. Sign languages are signed and seen. Spoken languages have been written on stone, wood, paper and electronic media. The technical support ranged from a chisel to a keyboard. The writing systems which developed have been under the influence of the particular language and the technical support. Having a hammer in your right and a chisel in the left makes it difficult to write from left to right. Having stable vowels motivates their representation in the written form. So how can sign languages be written for love letters, poems, verdicts and recipes?

One possible answer is SignWriting. SignWriting does not decompose a sign into phonemes, syllables or morphemes but body-parts, movements and face expressions and assigns a representation to each of them. Given such representations - e.g. an alphabet for potentially all sign languages - how may a keyboard, the input system, look like? How are the simple elements (body-parts, movements and face expressions) to be encoded in the computer and how the composed signs? As pictures, in Unicode or XML? How will this influence the input of signs, the layout and formatting of SignWriting documents, the possibilities to perform fuzzy matches on texts, in dictionaries, in the Internet? The papers written by Richard Gleaves, Valerie Sutton (*Signwriter*), Antônio Carlos da Rocha Costa, Graçaliz Pereira Dimuro, Juliano Baldez de Freitas (*A Sign Matching Technique to Support Searches in Sign Language Texts*), Angel Herrero (*A Practical Writing System for Sign Languages*), Steven Aerts, Bart Braem, Katrien Van Mulders, Kristof De Weerdt (Searching SignWriting Signs), Daniel Thomas Ulrich Noelpp (*Development of a new 'SignWriter' Program*) discuss these and related questions.

SignWriting, however, is by no means the only possible way of writing signs. Thomas Hanke in his invited talk "HamNoSys – *Representing Sign Language Data in Language Resources and Language Processing Contexts*" introduces an alternative approach, the Hamburg Notation System for Sign Languages. The purpose of HamNoSys has never been a usage in everyday communication. It was designed to comply with research requirements, e.g. for corpus annotation, sign generation, machine translation and dictionary construction. It thus differs from SignWriting in its scope and granularity. Unicode and XML solutions are available for HamNoSys, c.f. Ralph Elliott, John Glauert, Vince Jennings and Richard Kennaway in their contribution "*An Overview of the SiGML Notation and SiGMLSigning Software System*".

Once these fundamental questions regarding the writing of sign languages will be settled, derived notions such as word n-grams and character n-grams, important for computational approaches, may be used for applications such as language recognition, document classification and information retrieval. Spelling checking, syntax checking and parsing are obvious further developments once these more fundamental questions about the writing of signs will have been agreed upon.

It is a matter of fact, however, that most signers have not been trained in reading or writing in SignWriting. What is known as "text-to-speech" in the processing of spoken languages would seem a possible solution: a front-end to web-pages, mail boxes etc. would sign out the written text. As shown by Maria Papadogiorgaki, Nikos Grammalidis, Nikos Sarris, Michael G. Strintzis in "*Synthesis of virtual Reality Animations from SWML using MPEG-4 Body Animation Parameters*" and Yiqiang Chen, Wen Gao, Changshui Yang, Dalong Jiang and Cunbao Ge in "*Chinese Sign Language Synthesis and Its Applications*", avatars, i.e. virtual signers, may be constructed which translate a written form of a sign language or spoken language into signs, just like translating "d" into the corresponding sound wave.

A front-end on the input side of the system might translate signs into a written representation. Speech Recognition becomes Sign Recognition. Two different techniques are introduced. The recognition with the help of a data glove precedes from the signer's perspective and his/her articulations, c.f. Jose L. Hernandez-Rebollar's contribution "*Phonetic Model for Automatic*

Recognition of Hand Gestures". This approach may seem in line with the definition of phonemes in terms of their articulation and not their acoustic properties. On the other hand, it does not match our every-day experience in which we use a microphone and not electronic contact points at our vocal cords, tongue, velum, teeth and lips when using a telephone. The recognition of signs with the help of cameras, the second alternative, leads to the description of signs from the observer's point of view, in terms of formants and f_0, so to say. However, the articulation can be reconstructed and might be a better representation for the signs than the 'phonetic' description, as suggested by Boris Lenseigne, Frédérik Gianni, and Patrice Dalle in "*A New Gesture Representation for Sign Language Analysis*".

Both modules, sign recognition and sign generation, may serve MT systems with a sign language as source or target language respectively. A sign language as target language is used in translation experiments described by Jan Bungeroth and Hermann Ney in "*Statistical Sign Language Translation*". This corpus-based approach to Machine Translation, by the way, raises the question of sign language corpora. The only paper which really tackles the question of signed corpora in this collection is that of Onno Crasborn, Els van der Kooij, Daan Broeder, Hennie Brugman "*Sharing sing language corpora online. Proposals for transcription and metadata*". Matt Huenerfauth in his contribution "*Spatial Representations for Generating Classifiers Predicates in an English to American Sign Language Machine Translation System*", focuses on a particularly difficult aspect of sign language generation, the classifier predicates. Thus, when signing "leaves are falling", it is not enough to generate the sign "leave" and "falling", e.g. a downward movement. Instead the hand shape of "falling" should indicate the kind of object that is falling, e.g. with a flat hand.

The usage of classifiers leads us directly to the question of how to construct dictionaries for sign languages. Learners' dictionaries, reference dictionaries, dictionaries of NLP applications all need information about part of speech, lexical functions, idioms, subcategorization and semantics, which by no means is the same as in the national spoken language.
How do we search in a sign language dictionary? Have you ever looked up a Chinese or Japanese Dictionary? Paola Laterza and Claudio Baj in their paper "*Progetto e-LIS@*" propose an at least partially equivalent approach to the ordering of signs in a sign language dictionary.
How do you present the dictionary content to a learner? In the national spoken language or in SignWriting? The complexity of the question can be gauged from Elana Ochse's contribution "*A Language via Two Others, Learning English through LIS*". Should we use videos, photos, animations or drawings to represent the entries in dictionaries? A number of authors discuss these and related topics in the context of specific dictionary projects: for static presentations, i.e. paper dictionaries, Inge Zwitserlood and Doeko Hekstra propose the "*Sign Printing System – SignPS*" to compose pictures of signs. Eleni Efthimiou, Anna Vacalopoulou, Stavroula-Evita Ftinea, Gregory Steinhauer focus in their paper "*Multipurpose Design and Creation of GSL Dictionaries*" on the content, i.e. the types of information to be included in a sign language dictionary. Chiara Vettori, Oliver Streiter and Judith Knapp focus on different user requirements and the possible role of SignWriting in a sign language dictionary. Rubén Nogueira, Jose M. Martínez and present a dictionary project of a particular kind: "*19th Century Signs in the Online Spanish Sign Language Library: the Historical Dictionary Project*." Ingvild Roald finally gives a practical account on the history of techniques for the creation of sign language dictionaries, discussing advantages and drawbacks of the respective approaches.

When writing these lines, the preparation of the workshop and the proceedings is almost finished. This workshop wouldn't have been possible without the energy many people have invested in their spare time. First of all we would like to thank the authors who have done their best and provided superb papers. Our thank goes also to the reviewers for their detailed and inspiring reviews. Last but not least we want to thank Sara Goggi who accompanied the workshop on behalf of the LREC Programme Committee.

In closing we would like to thank you for attending the workshop, and we wish you will have a great time.

Oliver Streiter and Antônio Carlos da Rocha Costa
April 22, 2004

HamNoSys – Representing Sign Language Data in Language Resources and Language Processing Contexts

Thomas Hanke

University of Hamburg
Institute of German Sign Language and Communication of the Deaf
Binderstraße 34, 20146 Hamburg, Germany
Thomas.Hanke@sign-lang.uni-hamburg.de

Abstract

This paper gives a short overview of the Hamburg Notation System for Sign Languages (HamNoSys) and describes its application areas in language resources for sign languages and in sign language processing.

1. Introduction

The Hamburg Notation System for Sign Languages (HamNoSys) is an alphabetic system describing signs on a mostly phonetic level. As many sign notation systems developed in the last 30 years, it has its roots in the Stokoe notation system that introduced an alphabetic system to describe the sublexical parameters location, hand configuration (in most cases, the handshape only) and movement to give a phonological description of American Sign Language signs (Stokoe, 1960).

HamNoSys (first version defined in 1984, first published version Prillwitz et al., 1987), however, was designed to be usable in a variety of contexts with the following goals in mind:

- International use: HamNoSys transcriptions should be possible for virtually all sign languages in the world, and the notation should not rely on conventions differing from country to country, such as the national fingerspelling alphabets.
- Iconicity: As the large number of possible parameter variations did not allow for a standard alphabet (e.g. Roman alphabet) familiar to the users, newly created glyphs should be designed a way that helps to memorise or even deduct the meaning of the symbols wherever possible.
- Economy: While it should be possible to transcribe any signed utterance (even sign errors) with HamNoSys, notation of the majority of signs should make use of principles such as symmetry conditions, resulting in much shorter notation for the average sign.
- Integration with standard computer tools: The notation system should be usable within computer-supported transcription as well as in standard text processing and database applications.
- Formal syntax: The notation language should have a well-defined syntax, and its semantics should generally follow the compositionality principle.
- Extensibility: As it seemed obvious that, given the state of the art in sign language research, a notation system would not be capable of addressing all aspects of sign formation description for all sign languages right from the beginning, HamNoSys should allow both for a general evolution and specialisations. New versions of the system should not render old transcriptions invalid.

More than fifteen years after the first published version, HamNoSys is now at version 4 (Schmaling/Hanke, 2001). This latest version filled some minor gaps and introduced some shortcuts, but more importantly addressed issues related to using HamNoSys in a sign language generation context. For the latter purpose, it was also complemented with a new set of systems to encode nonmanual behaviour in a detailedness not previously possible in HamNoSys.

2. Overview of the System

2.1. General Structure

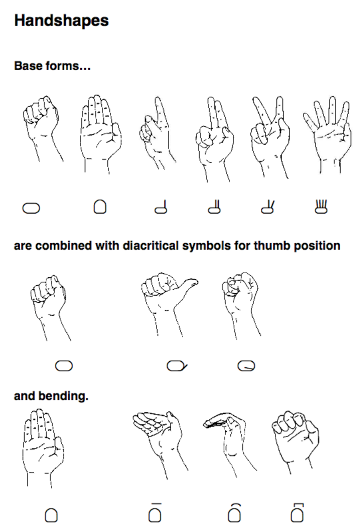
A HamNoSys notation for a single sign consists of a description of the initial posture (describing nonmanual features, handshape, hand orientation and location) plus the actions changing this posture in sequence or in parallel. For two-handed signs, the initial posture notation is preceded by a symmetry operator that defines how the description of the dominant hand copies to the non-dominant hand unless otherwise specified.

Specifications of nonmanual features and actions are optional. If the location specification is missing, a default location is assumed.

2.2. Handshapes

The description of a handshape is composed of symbols for basic forms and diacritics for thumb position and bending. In addition, deviations from this general description with respect to the fingers involved or the form of individual fingers can be specified. Where necessary, intermediate forms can be described as well.

By this combinatorial approach, the set of describable handshapes is rather large and is

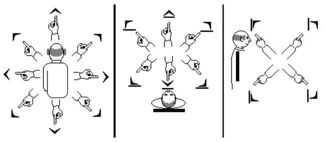


supposed to include all handshapes actually used in sign languages documented so far.

Dynamic handshapes as defined for German Sign Language by Prillwitz et al. (2002) are not considered primitives in HamNoSys. Instead, the initial handshape of an opening or closing dynamic pair appears within the posture, whereas the second one appears as the target of a handshape change action. For wiggling etc., one representative handshape is described in the posture, the wiggling itself, however, is described as an action.

2.3. Hand Orientation

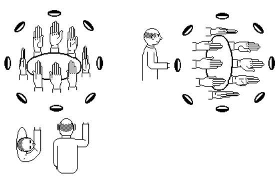
HamNoSys describes the orientation of the hand by combining two components: extended finger direction (i.e. for index hands the index direction) specifying two degrees of freedom, and palm orientation determining the third degree. By providing symbols for both components in a distance of 45°, a sufficiently fine-grained determination of the 3D-orientation of the hand becomes possible.



The three perspectives used for the extended finger direction (signer's view, birds' view, and view from the right) are reflected in the glyphs by no reference line, a horizontal reference line, or a vertical reference line representing the signer's body. (The same model is used for movements.)

Redundant symbols, such as , are not used. Insofar, there is a priority ordering between the three views determining which view to be used for each symbol.

For the third degree of freedom, only eight symbols are needed. The meaning of a symbol is defined relative to the extended finger direction (palm down, palm away from the body etc.).

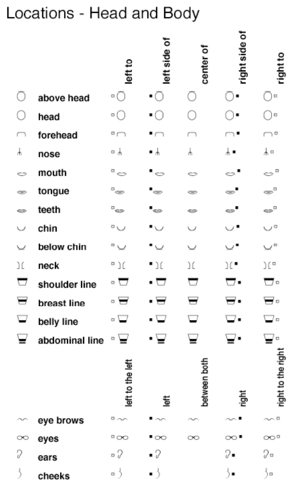


By adding a subscript, hand orientation can be made relative to the movement, e.g. the palm orientation changes as the movement direction changes:

HOUSE

2.4. Location

As with hand orientation, location specifications are split into two components: The first determines the location within the frontal plane (x and y coordinates, whereas the second determines the z coordinate. If the second part is missing, a "natural" distance of the hand from the body is assumed. If both parts are missing, the hand is assumed to be located in "neutral signing space", i.e. with "natural" distance in front of the upper torso.

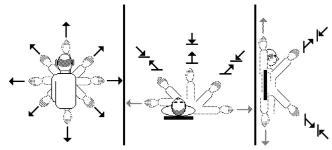


For two-handed signs, the location may also describe the relation of the two hands to each other ("hand constellation") as describing the positions of the two hands with respect two body parts might not be precise enough.

2.5. Actions

Actions are combinations of path movements (i.e. movements changing the position of the hand) and in-place movements of the hands as well as nonmanual movements. The combinations can be performed either sequentially or cotemporally.

In HamNoSys, path movement buildings blocks are straight lines, curved and zigzag lines, circles and similar forms. Here again, a quantization with 45° is applied.



Path movements can be specified either as targeted movements (target specified as location) or relative movements (target determined by the direction and the size of the movement).

In-place movements are changes in handshape or hand orientation as well as wiggling, twisting etc.

For all movement components, diacritic symbols to specify size can be added. Furthermore, for each movement a mode (such as slow or sudden stop) can be specified.

Repetitions of actions can be specified either by exact numbers as multiple repetition. In each case, a repetition can be continuous or recurrent.

The mere concatenation of actions means their performance in sequence, whereas actions in square brackets are done in parallel. E.g. a circle movement in square brackets with a straight movement results in a spiral movement. For two-handed actions, it is possible to specify different actions for each hand to be performed simultaneously.

2.6. Two-handed Signs

The notation of a two-handed sign begins with a symmetry marker. This symbol determines how to copy the specification for the dominant hand to the non-dominant hand. Exceptions can always be specified by separately describing configurations or actions for each hand. Example:

(German Sign Language NINETEEN): Both hands have the same hand orientation and the same movement, but they differ in their handshapes.

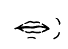
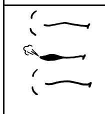
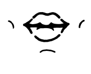
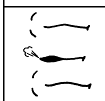
2.7. Nonmanual Components

As most notation systems, HamNoSys focuses on the description of the manual activities within a sign. The descriptive power of the existing system with respect to nonmanuals is rather limited: For each action, HamNoSys allows to specify an articulator to replace the hand. The actions available are those introduced for the hands. This allows appropriate descriptions for shoulder shrugging, head movements etc. but not necessarily facial expressions or mouth movements.

Originally, it was planned to add a facial circle to be complemented with diacritics for eyes, eyebrows, nose, cheeks, and mouth. At that time, however, practical limitations did not allow for the sheer number of diacritical symbols to be put into one font. Later suggestions added movement primitives to HamNoSys that targeted towards facial movements.

For the time being, we use a rather unsophisticated coding scheme to specify a number of nonmanual tiers in a multi-tier transcription scheme with the HamNoSys manual tier being the master tier. Synchronization is generally done on a sign level only.

Coding schemes are defined for eye gaze, facial expression (eye brows, eye lids, nose), mouth gestures and mouth pictures. The separation from the manual parts allows codes to be defined for states as well as for movements, i.e. sequences of states (e.g. TB tightly shut eyelids vs. BB eye blink). For mouth gestures, the coding scheme simply enumerates all gestures identified so far, e.g.:

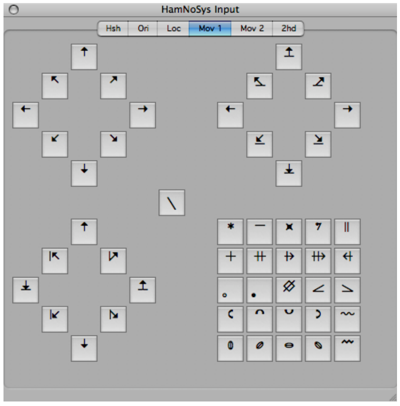
	C01	cheeks puffed	(static)
	C02	cheeks and upper and lower lip areas puffed	(static)
	C03	cheeks puffed gradually	(dynamic)
	C04(C)	one cheek puffed	(static)
	C05(C)	one cheek puffed; blow out air briefly at corner of one‘s mouth	(dynamic)
	C06(C)	one cheek puffed; blow out air briefly at corner of one’s mouth when touch-ing cheek with index finger	(dynamic)
	C07	cheeks sucked in, without sucking in air	(static)
	C08	cheeks sucked in, sucking in air through slightly open lips	(dynamic)
	C09(C)	tongue pushed into cheek (visible from outside)	(static)
	C10(C)	tongue pushed into cheek several times (visible from outside)	(dynamic)
	C11(C)	one cheek puffed; blow out air briefly at corner of one‘s mouth several times	(dynamic)
	C12	lips closed, tongue pushed behind bottom lip/chin (visible from outside)	(static)

A complete documentation of these nonmanual coding schemes can be found in Hanke et al. (2001).

2.8. Implementation

The HamNoSys symbols are available as a Unicode font, with the characters mapped into the Private Use area of Unicode.

For MacOS X, a keyboard layout has been defined that can be automatically activated once text in the HamNoSys font is selected. This keyboard graphically arranges the characters on the keyboard, e.g. the arrows in circles with 45° sectors. This makes learning keyboard input rather easy for those using HamNoSys every day. For people who use the system less frequently, even this keyboard is too much to memorise. Here we offer (for both MacOS and Win) a small input utility that allows the user to construct the HamNoSys string by clicking on the appropriate symbols on (user-configurable) palettes:



A syntax-oriented editor was available for HamNoSys 2 (Schulmeister, 1990), but has not been updated since then. Within the ViSiCAST project (cf. Schulmeister, 2001), SiGML, an XML equivalent to HamNoSys, has been defined (Elliott et al., 2000).

3. Dictionaries

In many sign language dictionaries, you find notation as a description how to perform an entry. Depending on the media used, the notation is part of a multitude of form descriptions, e.g. video, photos or drawings with or without arrows, performance instructions in written language, etc. Today’s sign language dictionaries mostly present only the citation form of the sign, some possibly add unstructured information like “directional verb” to indicate the kind and richness of morphological derivation that can be applied to the sign.

Notation is also used to provide some means of access to the dictionary contents from the sign language side: For search access, you normally find partial matching strategies. In the case of HamNoSys with its relatively high degree of detailedness, we add fuzzy search mechanisms to allow for variation. For browsing access (and this includes of course printed dictionaries), the lexemes (or an index thereof) are ordered according to only some parameters expressed in the notation. For HamNoSys, it is obvious why the order on HamNoSys strings induced by some order of the HamNoSys alphabet is not really useful: With about 200 symbols, no user will be able to memorise this ordering, and, for a given sign, you often find several equivalent HamNoSys notations, and HamNoSys still lacks a reduction method to identify one of these as the canonical notation. (For an example, cf. Konrad et al., 2003.)

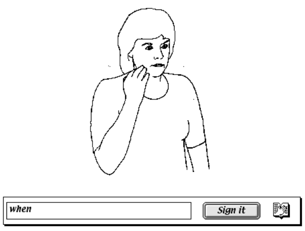
4. Transcription of Signed Corpora

Notation is used to transcribe linguistic data (by viewing video) to provide an efficient form description and to make it accessible to analysis. In the first years, notation was that part of the transcription that came closest to raw data. But even after the integration of digital video, notation did not become superfluous as it makes data searchable for phonetic aspects (cf. Hanke/Prillwitz, 1995 and Hanke, 2001).

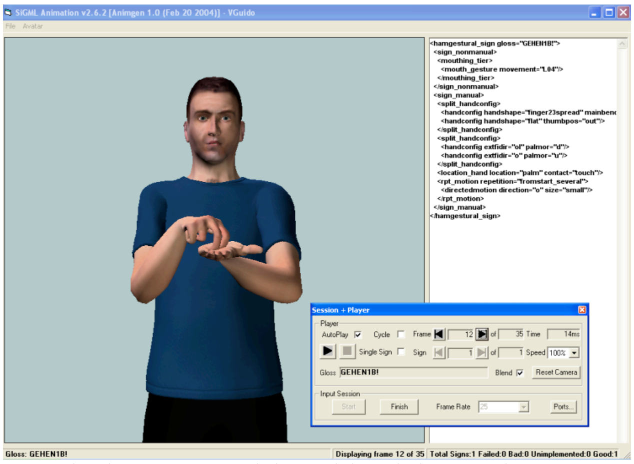
In most cases, the notation was used to describe the observed event at a certain level of detailedness. No attempt was made to relate the observed token to its type. One exception is the work by Johnston (1991), who, after giving the citation form of the type, describes how the token deviates from the type. In the context he introduced this notational convention, he considered those derivations only that are morphologically relevant, but it is easy to see how this could be extended.

ilex, our recent approach to corpus transcription (Hanke, 2002a), ties a lexicon into the transcription system and requires the user to relate each token to a type, a function considered absolutely necessary to ensure data consistency in larger corpora transcriptions that usually are team efforts and therefore cannot rely on the transcriber's intimate knowledge of the data already processed. What may be substituted in spoken language corpora by automatically searching the transcription data cannot be avoided for sign language corpora as long as HamNoSys or other notation systems do not establish a working orthography.

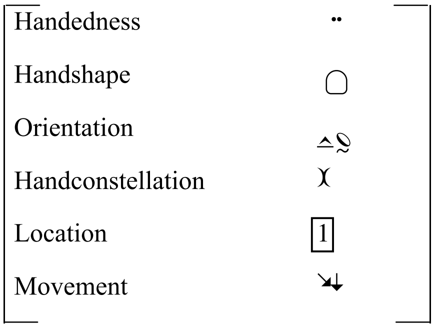
5. Generation



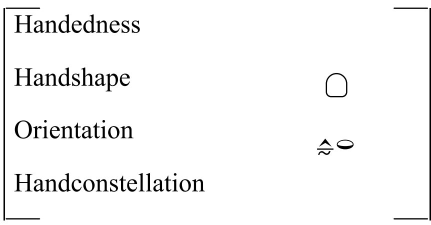
One of the first projects HamNoSys was used in is H.AN.D.S. (Hamburg Animated Dictionary of Signs, cf. Prillwitz/Schulmeister, 1987) which represented dictionary entries by the notation and a two-dimensional animation automatically created from the notation. Due to the immense number of high-precision drawings needed for that purpose, only a subset of HamNoSys could be correctly animated at the end of the project. The upcoming digital video technology then pushed animation to the background as far as sign language dictionaries were concerned. However, in the context of spoken-to-sign language translation systems, animation promises far better results than digital video: While Krapez/Solina (1999) describe a method to improve sign-to-sign video blending, they also outline the limitations. Animation technology can not only model transitional movements between signs, but, based on a suitable language model, provide uncountable morphological variations of sign lexicon entries as needed for locatable signs, directional verbs etc. Kennaway (2002, 2004) describes the ViSiCAST animation component based on SiGML:



The language model used in ViSiCAST is an HPSG feature structure. Depending on the morphological richness of a lexical entry, the structure may be fully instantiated with HamNoSys values, or might contain more complex structures only finally reducible into HamNoSys values. For a locatable sign like HOUSE in German Sign Language, this roughly looks as follows:



Using HamNoSys symbols as HPSG feature values is quite convenient as the user can immediately grasp the meaning of the values, and the approach has been successfully applied to a range of sign language specific phenomena such as classifier and number incorporation, directional verb signs and locatable signs. Problems remain where the independence of sign parameters is an over-simplification. This is easily illustrated with the example MOVE–classifier:car–source:right-side-in-front-of-the-signer–goal:left-side-in-front-of-the-signer. Once the feature structure for double-track vehicles



is unified with the lexical entry for MOVE and everything from source and goal except the height in signing space, the result is equivalent to the following lambda expression:

With heights above chest level, this results in highly unnatural signing: Instead of

one would sign

Apparently the assumption that a classifier feature structure should specify whole handshapes and hand orientations is too restrictive. Instead, one might want to specify a part of the hand and this part's orientation. While it is always possible to translate the fully instantiated structures into standard HamNoSys to feed into the animation component, this would distribute the need for anatomical knowledge over two system components: The language model and the animation component, a highly undesirable situation. Instead, it might be a better idea to allow parts of handshapes and orientations thereof instead of complete handshapes with hand orientation in HamNoSys itself. A suggestion in this direction also discussing other classes of examples has been made by Hanke (2002b).

6. Applications beyond Sign Language

While Miller (2001) reports that HamNoSys and the family of derivatives of the Stokoe notation are the most widely used systems in research, it seems that even more people use HamNoSys outside sign language research, namely in gesture research.

In the Berlin Dictionary of Everyday Gestures (Posner et al., in prep.), HamNoSys is used in toto to describe the form of the gestures in addition to photos and verbal descriptions.

A number of gesture description schemes inherit structure and/or feature values from HamNoSys, such as MURML (Kopp et al., 2004), FORM (Martell, 2002) and CoGesT (Gut et al., 2003). KINOTE (Hofmann/Hommel, 1997) was described by the authors as a kinematic transcription of a subset of HamNoSys. Some of these representation languages are also the target for gesture recognition, be they based on data gloves or video, so that HamNoSys is indirectly also used in recognition contexts for gesture.

7. Outlook

New application areas will always pose new requirements on a system such as HamNoSys. So we currently see no end in the development of the system.

Obviously, one application area for HamNoSys is still missing: Sign language recognition. Only a few sign language recognition systems work on a sublexical level (e.g. Vogler/Metaxas, 2004), and all recognition systems today work with rather small sign inventories. In the future, language models in connection with lexicons might help recognition systems to cover larger subsets of a sign language, and it would be interesting to see how HamNoSys fits into such a system.

For transcription schemes of signed texts as well as lexicon building, data on intra- and inter-transcriber reliability could contribute to another aspect to the question how useful a phonetic transcription of signs is.

The use of HamNoSys in a number of gesture description systems might turn out to be a useful key to link sign language resources and processing models with the larger field of multimodality.

8. References

Elliott, R., J. Glauert, R. Kennaway and I. Marshall, 2000. The development of language processing support for the ViSiCAST project. In *The fourth international ACM conference on Assistive technologies ASSETS*. New York: ACM. 101–108.

Gut, U., K. Looks, A. Thies, T. Trippel and D. Gibbon, 2003. CoGesT - Conversational Gesture Transcription System - version 1.0. http://www.spectrum.uni-bielefeld.de/modelex/publication/techdoc/cogest/CoGesT_1.pdf.

Hanke, T., 2001. Sign language transcription with syncWRITER. *Sign Language and Linguistics* 4(1/2): 267–275.

Hanke, T., 2002a. iLex - A tool for sign language lexicography and corpus analysis. In M. González Rodriguez and C. Paz Suarez Araujo (eds.), *Proceedings of the third International Conference on Language Resources and Evaluation. Las Palmas de Gran Canaria, Spain.* Paris: ELRA. 923-926.

Hanke, T., 2002b. HamNoSys in a sign language generation context. In R. Schulmeister and H. Reinitzer (eds.), *Progress in sign language research: in honor of Siegmund Prillwitz / Fortschritte in der Gebärdensprachforschung: Festschrift für Siegmund Prillwitz*. Seedorf: Signum. 249–266.

Hanke, T. and S. Prillwitz, 1995. syncWRITER: Integrating video into the transcription and analysis of sign language. In H. Bos and T. Schermer (eds.), *Sign Language Research* 1994*: Proceedings of the Fourth European Congress on Sign Language Research, Munich, September 1-3, 1994.* Hamburg: Signum. 303–312.

Hanke, T., G. Langer and C. Metzger, 2001. Encoding non-manual aspects of sign language. In: T. Hanke (ed.), *Interface definitions*. ViSiCAST Deliverable D5-1.

Hofmann, F. and G. Hommel, 1997. Analyzing human gestural motions using acceleration sensors. In P. Harling and A. Edwards (eds.), *Progress in gestural interaction. Proceedings of Gesture Workshop '96.* Berlin: Springer. 39-59.

Johnston, T., 1991. Transcription and glossing of sign language texts: Examples from AUSLAN (Australian Sign Language). *International Journal of Sign Linguistics* 2(1): 3–28.

Kennaway, R., 2002. Synthetic animation of deaf signing gestures. In I. Wachsmuth and T. Sowa (eds.), *Gesture and sign language in human-computer interaction. International Gesture Workshop, GW 2001, London, UK, April 18-20, 2001 Proceedings*. Berlin: Springer. 146-157.

Kennaway, R., 2004. Experience with and requirements for a gesture description language for synthetic animation. In A. Camurri and G. Volpe (eds.), *Gesture-Based Communication in Human-Computer Interaction, 5th International Gesture Workshop, GW 2003, Genova, Italy, April 15-17, 2003, Selected Revised Papers.* Berlin: Springer. 300-311.

Konrad, R. et al., 2003. *Fachgebärdenlexikon Sozialarbeit/Sozialpädagogik.* Hamburg: Signum 2003.

[WWW version at http://www.sign-lang.uni-hamburg.de/solex/]

Kopp, S., T. Sowa and I. Wachsmuth, 2004. Imitation games with an artificial agent: from mimicking to understanding shape-related iconic gestures. In A. Camurri and G. Volpe (eds.), *Gesture-Based Communication in Human-Computer Interaction, 5th International Gesture Workshop, GW 2003, Genova, Italy, April 15-17, 2003, Selected Revised Papers.* Berlin: Springer. 436-447.

Krapez, S. and F. Solina, 1999. Synthesis of the sign language of the deaf from the sign video clips. *Elektrotehniski vestnik.* 66(4-5). 260-265.

Martell, C., 2002. Form: An extensible, kinematically-based gesture annotation scheme. In M. González Rodriguez and C. Paz Suarez Araujo (eds.), *Proceedings of the third International Conference on Language Resources and Evaluation. Las Palmas de Gran Canaria, Spain.* Paris: ELRA. 183-187.

Miller, C., 2001. Some reflections on the need for a common sign notation. *Sign Language and Linguistics.* 4(1/2):11-28.

Posner, R., R. Krüger, T. Noll and M. Serenari, in prep. *Berliner Lexikon der Alltagsgesten*. Berlin: Berliner Wissenchafts-Verl.

Prillwitz, S. and R. Schulmeister, 1987. Entwicklung eines computergesteuerten Gebärdenlexikons mit bewegten Bildern. *Das Zeichen* 1(1): 52-57.

Prillwitz, S. et al., 1987. *HamNoSys. Hamburg Notation System for Sign Languages. An introduction*. Hamburg: Zentrum für Deutsche Gebärdensprache.

Prillwitz, S. et al., 2002. Zur Phonetik und Phonologie der Deutschen Gebärdensprache. Manuscript, University of Hamburg.

Schmaling, C. and T. Hanke, 2001. HamNoSys 4.0. In: T. Hanke (ed.), *Interface definitions*. ViSiCAST Deliverable D5-1. [This chapter is available at http://www.sign-lang.uni-hamburg.de/projekte/HamNoSys/HNS4.0/englisch/HNS4.pdf]

Schulmeister, R., 1990. Zentrum erhält den Deutschen Hochschul-Software Preis 1990. *Das Zeichen* 4(11): 73-77.

Schulmeister, R., 2001. The ViSiCAST Project: Translation into sign language generation of sign language by virtual humans (avatars) in television, WWW and face-to-face transactions. In C. Stephanidis (ed.), *Universal access in HCI : towards an information society for all.* Hillsdale, NJ: Erlbaum. 431-435.

Stokoe, W., 1960. *Sign language structure: An outline of the visual communication systems of the American deaf.* Buffalo, NY: Univ. of Buffalo.

Vogler, C. and D. Metaxas, 2004. Handshapes and movements: multiple-channel American Sign Language recognition. In A. Camurri and G. Volpe (eds.), *Gesture-Based Communication in Human-Computer Interaction, 5th International Gesture Workshop, GW 2003, Genova, Italy, April 15-17, 2003, Selected Revised Papers.* Berlin: Springer. 247-258.

SIGNWRITER

Richard Gleaves
Valerie Sutton

Deaf Action Committee for SignWriting
Center For Sutton Movement Writing
La Jolla, California, USA
rgleaves@signwriting.org
sutton@signwriting.org

Abstract

This paper reviews the design history of SignWriter, a word processor for the SignWriting system. While the primary goal of SignWriter was simply to create a word processor for SignWriting, its development and subsequent use had several beneficial effects on the SignWriting system. Various design aspects of SignWriter are considered in the context of current computing technologies and sign processing development efforts.

Background

The SignWriting system [Sutton04] was conceived, developed, and used for many years as a hand-written notation. In particular, its use predated the introduction of low-cost personal computers.

In 1984 Emerson and Stern Associates, a small educational research and development firm, received a grant to develop a word processor for SignWriting. The resulting software, which operated on an Apple II computer, supported only a minor subset of the SignWriting system and was more of a demonstration than a useful tool: it was not subsequently used, and received no further development. The application was notable for displaying the symbols in a virtual "picture frame" around a central editing area, with symbols selected for entry by moving a cursor around the frame until the desired symbol was reached.

Emerson and Stern's software design implied that SignWriting was too complex for the personal computers of the time. Interestingly, their response was to devise an entirely different writing system named SignFont [Newkirk87], which traded computational simplicity - it was designed as a standard Macintosh font - for notational obscurity. SignFont's subsequent nonuse suggests that this design tradeoff was unsuccessful.

SignWriter Apple

It was in this context that SignWriter was conceived in 1986. The intended use for SignWriter was in education and the hardware platform was once again the Apple II, which at the time was an established standard for personal computing. The design goal was to implement the full SignWriting system in a simple but complete and usable word processor.

This more ambitious goal could be attempted on the same hardware because as a former member of the UCSD Pascal project, Richard Gleaves had several years of experience developing system software for the Apple II, and knew how to program in assembly language and make full use of the Apple's 128KB memory. In addition Gleaves' Pascal project colleague Mark Allen provided some high-performance graphics routines that he had developed for writing arcade-style games on the Apple II.

Much of the design effort in SignWriter was spent on two issues:

… Developing a memory-efficient encoding for SignWriting text
… Devising user interface mechanisms for efficiently typing symbols

SignWriting symbols were encoded using a variable-length byte-code system that was introduced in UCSD Pascal p-code [Bowles78] and later adopted for use in Java object code. The SignWriter graphics engine interpreted the byte codes as instructions for drawing symbols on the screen in specific locations and orientations.

Typing was chosen as the input mode for two reasons. First, while mice were available for the Apple II they were an optional add-on and therefore most Apple IIs did not have them. Second, the SignWriting system was receiving criticism at the time for allegedly being a form of illustration rather than a true writing system. Therefore an efficient typing mechanism would cause SignWriter to serve as implicit proof that SignWriting was indeed a form of writing.

It was evident that SignWriting's complex symbol set would prevent it from being typed as efficiently as the Roman alphabet on a standard keyboard. However, the design that evolved - which involved the context-sensitive dynamic redefinition of the keyboard keys - yielded a valuable tradeoff of efficiency for learnability. The key boxes displayed on the screen highlighted the natural categories of the SignWriting symbols in a manner that allowed the typing mechanism to serve as an implicit learning tool: a crucial property given the symbol set complexity and the application's intended audience. See Figures 1, 2 and 3 from the SignWriter-At-A-Glance Instruction Manual.

The SignWriting symbol images were created by Valerie Sutton using the SignWriter symbol editor program. In addition she defined the mapping of SignWriting symbols to the keyboard keys. As with the key boxes, this mapping emphasized learnability by grouping symbols according to their natural categories. Conversely, the mapping of the key box keys and symbol attribute keys (Arrow, Cursor, Mirror, Size, and Rotate) was determined strictly by typing efficiency.

SignWriter's Find and Replace commands were implemented (at significant expense in memory) both to establish SignWriter as a complete word processor and again to demonstrate SignWriting's status as a true writing system. Unfortunately the search algorithm did not take into account the relative positioning of symbols within a sign, thus making the search feature itself more of a demonstration than a useful tool.

Because SignWriter was developed as a stand-alone application, it was free to possess an application-specific user interface. The interface design was influenced by Tufte's principle of graphical minimalism [Tufte83]: namely, every pixel that was not part of a SignWriting symbol existed onscreen only because it was functionally necessary. While this design approach may seem austere given today's large color displays, it made for a simple and easy-to-use interface on the Apple II, which had a screen resolution of only 560 by 192 pixels.

The major drawbacks to SignWriter's interface design were the inefficient cursor movement commands and the need for a keyboard card showing the assignment of SignWriting symbols and commands to the keys.

The Apple II version of SignWriter supported the full SignWriting system as it was defined at the time (palm orientation had not yet been introduced). The software was quite usable, but was never widely used because experienced SignWriting users had to type in each occurrence of each sign, while for new users typing symbols was relatively inefficient and – in the absence of a system for teaching typing – posed a significant learning curve.

SignWriter DOS

By the late 1980s the IBM PC had replaced the Apple II as the personal computer of choice. SignWriter was ported to the IBM PC with programming assistance from Barry Demchak. We chose the CGA display mode because at the time it was the graphics display mode supported by the most PC models, and because its screen resolution of 640 by 200 pixels was close enough to the Apple to simplify porting the existing symbol graphics to the PC (which is why the SignWriter symbols are so jagged).

The extra memory available on the IBM PC allowed SignWriter to be expanded with additional symbols, a sign dictionary, and support for multiple countries and languages. These features (along with software distribution on the Internet) had a significant impact on SignWriter use, as researchers began using SignWriter to create and publish dictionaries for various signed languages. This is the version of SignWriter that is in common use today.

Effects on SignWriting

The purpose of SignWriter was simply to provide a word processor for the SignWriting system. However, its development and subsequent use had several beneficial effects on SignWriting:

… SignWriter offered a concrete proof of SignWriting's status as a systematic notation rather than an ad hoc form of illustration. This notion influenced the subsequent design of the software.

… The typing mechanism served as an implicit interactive system for learning the SignWriter symbols (an important achievement given the complexity of the symbol set).

… The SignWriter symbol editor was withheld from distribution to ensure the controlled development of the

SignWriting system as it evolved to support more and more signed languages.
… The constraints of computer implementation exerted a positive influence on the subsequent evolution of the SignWriting system.
… The SignWriter software itself served as an efficient means of distributing the SignWriting system, and established a de facto standard for data exchange (an effect greatly amplified by the introduction of the Internet).

Conclusion

Beyond its immediate value as a tool for practical sign processing, SignWriter offers a number of lessons for current and future developers of sign processing software.

The most important is the need to standardize a user interface mechanism for symbol input; just as the symbol set is being standardized across all sign processing programs that use SignWriting, so must symbol entry. Such a standard should be centered on typing, with mouse input as an alternative rather than a replacement. Compelling pedagogical and linguistic reasons exist for providing efficient input mechanisms at the level of symbols rather than signs; while such mechanisms need not supplant text entry at the sign level, the reverse equally holds true.

The diagrams in this paper illustrate SignWriter's typing-based symbol input system as an example of how future typing-centered systems could be designed.

With regards to efficiency, Valerie Sutton has learned to type SignWriting almost as efficiently as English. This suggests that with the proper training (an accepted norm for typing) and appropriate hardware (e.g., a notebook computer with an integrated touchpad for cursor control and fine symbol positioning), typing-centered symbol input may well prove superior to any mouse-based systems.

Finally, SignWriter demonstrated that with the appropriate software architecture a true word processor could be implemented for SignWriting given limited resources for memory, processing power, and display resolution. This in turn suggests opportunities for developing useful sign processing software on the emerging handheld computing platforms such as PDAs and cell phones.

References

[Bowles78]
Bowles, Kenneth L., "UCSD Pascal", Byte. 46 (May)

[Newkirk87]
Newkirk, Don, "SignFont Handbook", San Diego: Emerson and Stern Associates (1987)

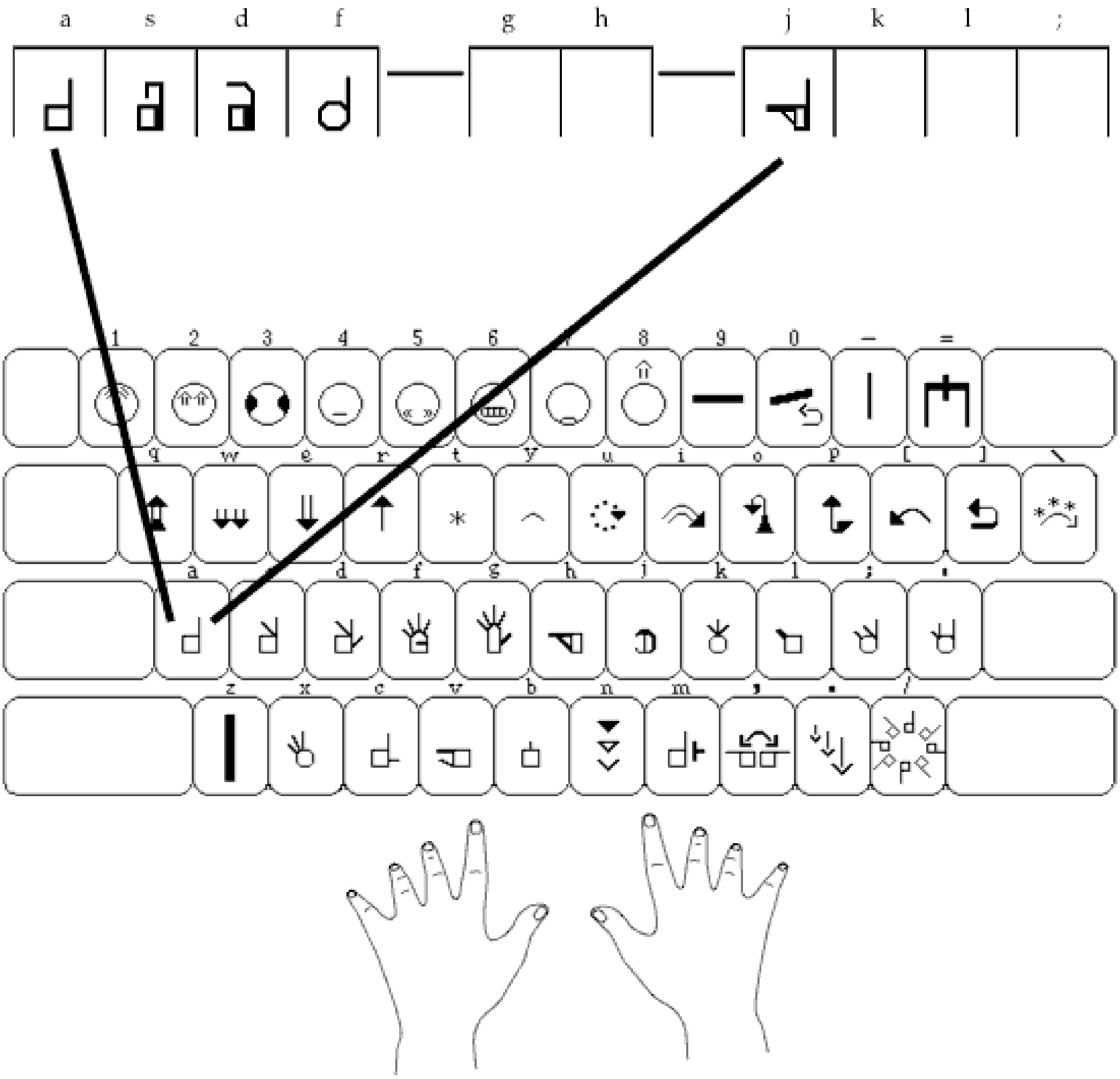
[Tufte83]
Tufte, Edward R., "The Visual Display of Quantitative Information", Graphics Press (1983)

[Sutton93]
Sutton, Valerie. SignWriter-At-A-Glance Instruction Manual, SignWriter Computer Program Notebook, Deaf Action Committee For SignWriting (1993)

[Sutton04]
Sutton, Valerie. SignWriting Site. www.signwriting.org

Typing Group 1

with the **SignWriter**® Computer Program



Group 1 is located on the **a** key.
When you press the **a** key, all of the hands in **Group 1** appear in a row on the screen.
You can then choose the hand symbol you wish, by pressing the **a, s, d, f, or j** keys.

Figure 1: A page from the SignWriter-At-A-Glance-Manual. Symbol groups are under each key.

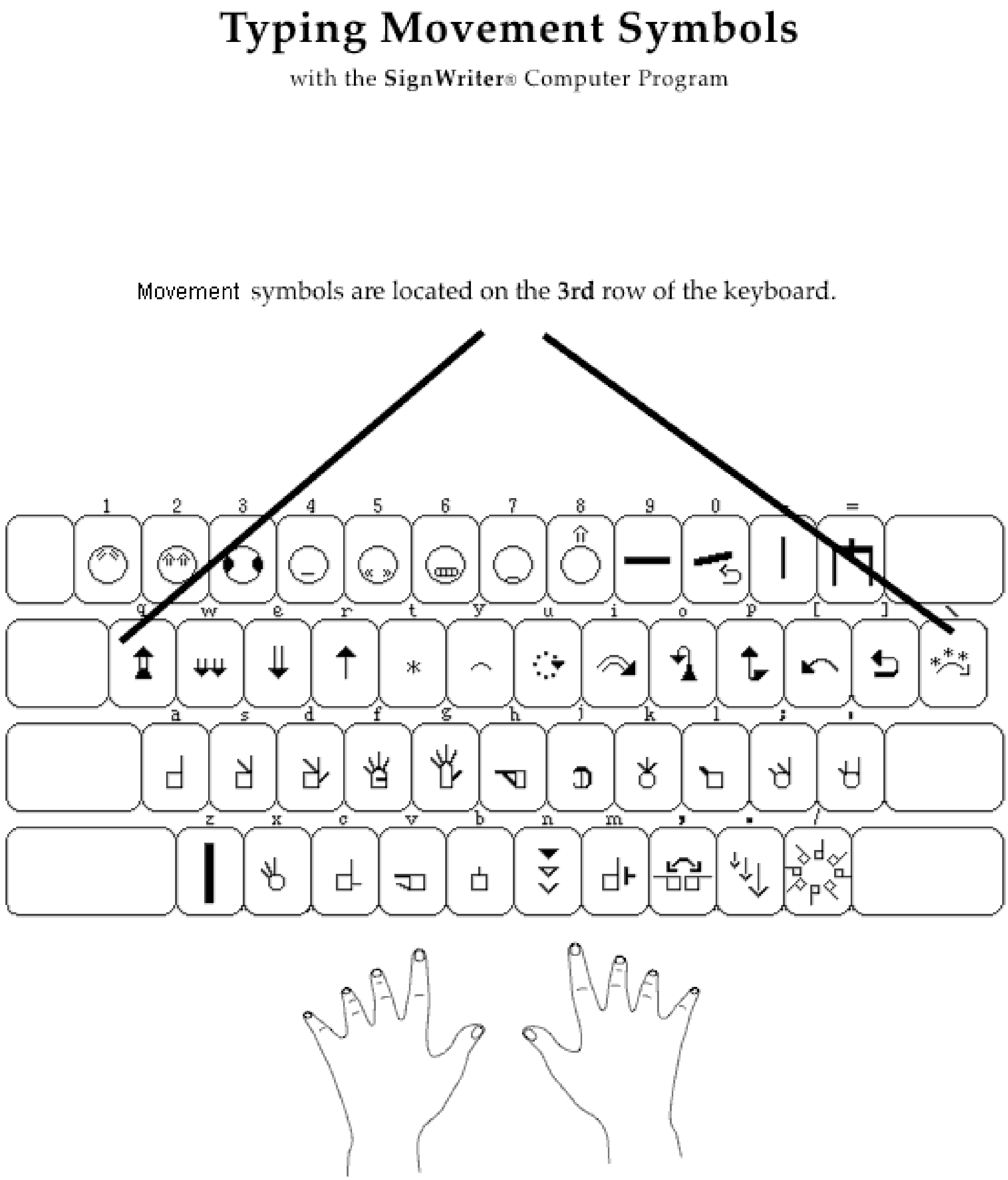
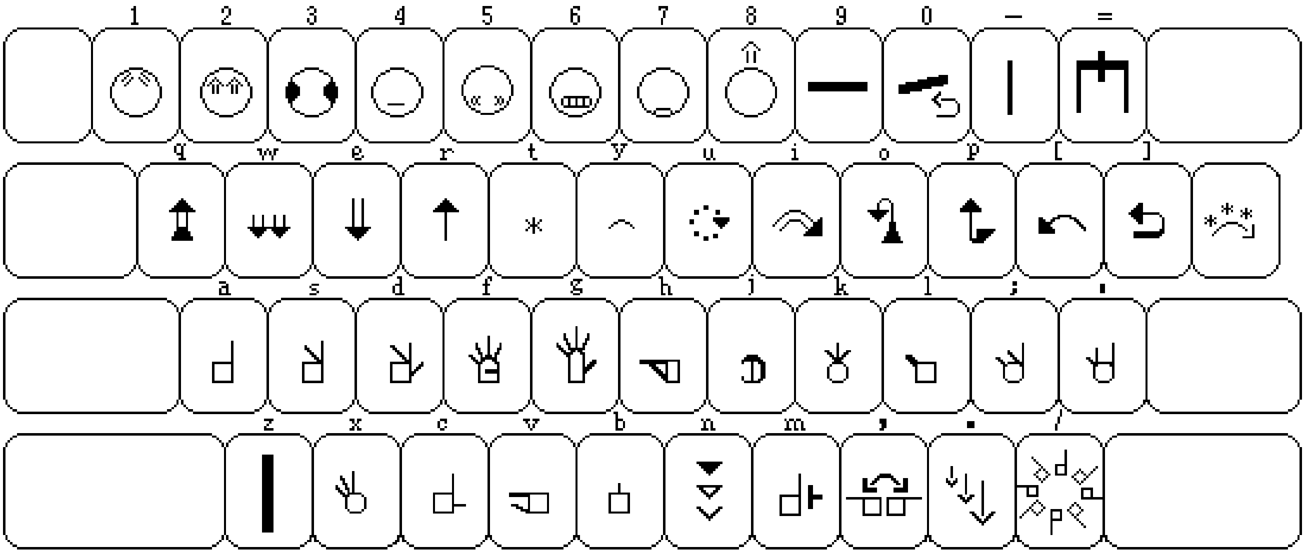
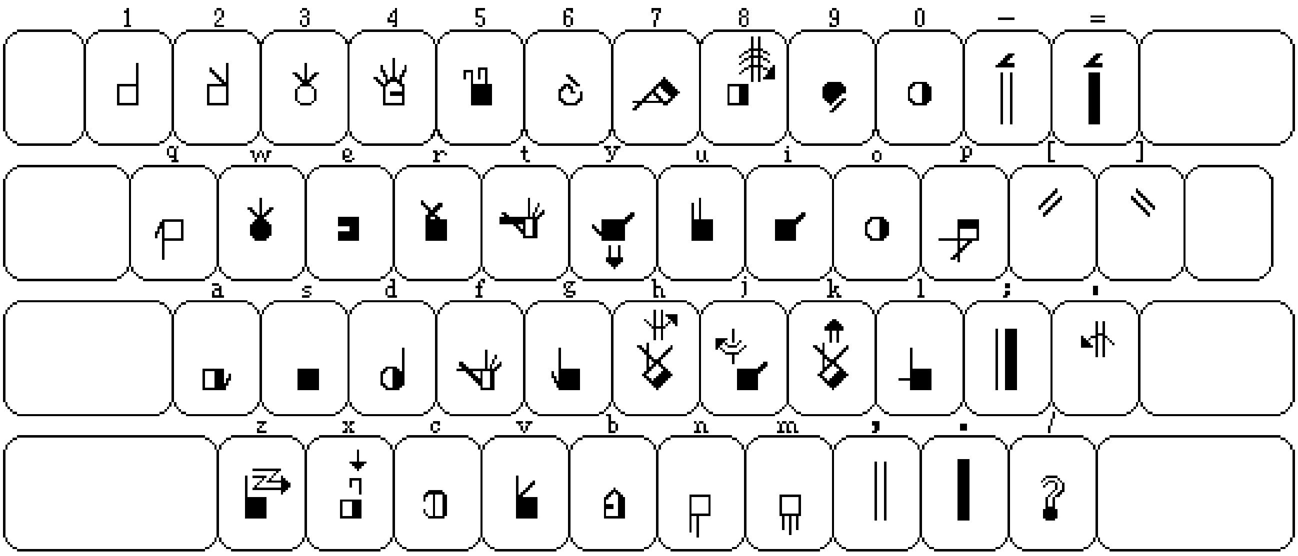


Figure 2: A page from the SignWriter-At-A-Glance-Manual. Symbol categories are placed in rows of keys.



Sign Keyboard
All Countries



Fingerspelling, Brazil
Country Code: 055

Figure 3: A page from the SignWriter-At-A-Glance-Manual. 17 countries with 17 fingerspelling keyboards.

Open-ended Resources in Greek Sign Language: Development of an e-Learning Platform

Galini Sapountzaki[1], Eleni Efthimiou[1], Kostas Karpouzis[2], Vassilis Kourbetis[3]

[1] ILSP-Institute for Language and Speech Processing
Artemidos 6 & Epidavrou, GR 151 25, Maroussi, Greece
[2] Image, Video and Multimedia Systems Laboratory, National Technical University of Athens, 9, Iroon Polytechniou Str., GR 157 73, Zographoy, Greece
[3] Pedagogical Institute, Messogion 392 Aghia Paraskevi GR 153 41, Athens, Greece
gsapountz@uth.gr, eleni_e@ilsp.gr, kkarpou@softlab.ece.ntua.gr,
kourbeti@pi-schools.gr

Abstract

In this paper we present the creation of dynamic linguistic resources of Greek Sign Language (GSL). The resources will feed the development of an educational multitask platform within the SYNENNOESE project for the teaching of and in GSL. The platform combines avatar and animation technologies for the production of sign sequences/streams, exploiting digital linguistic resources of both lexicon and grammar of GSL. In SYNENNOESE, the input is written Greek text, which is then transformed into GSL and appears animated on screen. A syntactic parser decodes the structural patterns of written Greek and matches them into equivalent patterns in GSL, which are then signed by a virtual human. The adopted notation system for the lexical database is HamNoSys (Hamburg Notation System). For the implementation of the digital signer tool, the signer's synthetic movement follows MPEG-4 standard and frame H-Anim with the use of VRML language.

1. Introduction

Primary target user group are the deaf pupils who need teaching tools and educational material for the GSL grammar class. Till very recently educational material was available to students with hearing impairments only in written Greek form. Formal teaching of GSL as a first language from the very early school years, and relevant development of educational content is becoming very urgent since law 2817/2000 was put into action by the Hellenic State. This law defines that «the official school language of deaf and hard 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 hearing students». In this context the new education programs of the Pedagogical Institute[1] (in print) 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.

In consultancy with the Pedagogical Institute, SYNENNOESE helps pupils acquire the proper linguistic background so that they can take full advantage of the new accessible 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.

2. Greek Sign Language – the background

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; some work has been done on individual aspects of its syntax (negation (Antzakas & Woll, 2001), morphology (Lampropoulou, 1992)), 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 (Lampropoulou, 1997). 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 analysed sign languages (Bellugi & Fischer, 1972 ; Liddell, 1980).

GSL has developed in a social and linguistic context similar to most other sign languages (Kyle & Woll, 1985 ; Brennan, 1987). It is used widely in the Greek deaf community and the estimation for GSL users is about 40,600 (1986 syrvey of Gallaudet Univ.). 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 Greek Federation of the Deaf (GFD) show that, in the year 2003 about 300 people were registered for classes of GSL as a second language. The recent increase of mainstreamed deaf students in 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. Official settings where GSL is being used include 11 Deaf

[1] Pedagogical Institute (PI) is the official organisation that validates all educational programs of primary and secondary education in Greece.

clubs in Greek urban centers and a total of 14 Deaf primary, secondary and tertiary educational settings.

3. Linguistic research background in the area of sign languages

In Greece there have been some serious attempts of lexicography in the recent past (PROKLESE, a Dictionary of Computing Signs, NOEMA: a Multimedia Dictionary of GSL Basic Vocabulary and A Children's Dictionary of GSL) mainly for educational purposes (Kourbetis, 1999 ; Kourbetis & Efthimiou, 2003), but complete decoding of the language structure is not yet publicly available.

The linguistic part of the project is based on overall assumptions for the adequacy of signed languages as by Stokoe (1960, 1978), Woll & Kyle (1985), Valli & Lucas (1995), Sutton-Spence & Woll (1999), Neidle et al. (2000), Gee & Goodhart, (1985) among many. Greek sign language is analyzed to its linear and non-linear (simultaneous) components (Padden, 1989 ; Engberg – Pedersen, 1993). 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. Each sign in GSL is described as to its handshape, location, movement, orientation, number of hands and use of any obligatory non-manually articulated elements (referred to as nmf, i.e. mouth patterns, head and shoulder movements and other non-manual features), based on the Stokoe model (ibid).

In the project it was considered essential that the output is as close to native GSL as used in the Greek deaf community. In this respect, forms of 'signed Greek' or other manual codes for the teaching of Greek were excluded and the two languages (GSL and Greek) were treated as the first and second language respectively for the users of the platform, quite as other bilingual platforms may function outside the domain of special education.

4. The project's language resources

Implementation of both the tutoring and the summarization tools of the platform require collection of extensive electronic language resources for GSL as regards the lexicon and the structural rules of the language (Efthimiou et al., 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 & Katsoyannou, 2001 ; Efthimiou & Katsoyannou, 2002). 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. The project 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 Greek Federation of the Deaf 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 analysis of GSL as a natural language. 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 recognised by native signers as correct utterances of the language.

In this respect SYNENNOESE offers a great challenge for in-depth work on both directions, lexicography and linguistic analysis of GSL; for the first time research will go beyond a mere collection of glosses (Logiadis & Logiadis, 1985) and move further from many previous bilingual dictionaries of sign languages (Brien & Brennan, 1992)), into the domain of productive lexicon (Wilcox et al., 1994), i.e. the possibility of building new GSL glosses following known structural rules, and also challenge automatic translation in predictable environments, using an effective module/interface for the matching of structural patterns between the written input and the signed output of the platform. 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. From a linguistic point of view the resulting database of glosses, rules and tendencies of GSL will be a significant by-product of the project, of great value to future applications.

4.1 Grammar content definition

In the early implementation phase, the subsystem for the teaching of GSL grammar covers a restricted vocabulary and a core grammar capable of analysing a restricted number of main GSL grammatical phenomena, which might be argued that belong to signing universals:

The objective of the 18-month project is to transcribe the digitized avi files with GSL individual signs and store them in a retrievable database. This 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 are analyzed in this early stage, so that feedback from the technical team will determine further steps (Johnston & Schembri, 1999). Non-manual grammatical features (Boyes Braem & Sutton-Spence, 2001) and polymorphemic signs are acknowledged but not included in this stage. In the second stage longer sequential structures of signs will be considered (e.g. compound word-signs) and once individual signs are transcribed and stored in a database, additional tiers such as non-manual features can be added without technical difficulties.

At the stage of grammatical analysis international findings on sign language grammars, as well as the views of our deaf native user consultants are taken into account in order to verify findings. It is admitted that there is even more work to be done on the pragmatics of GSL and its relation with real-world situations (e.g. for the use of indexes or classifiers), and these are noted as future aims of the platform.

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 proper nouns, acronyms, terminology or general terms for which no specific sign exists. Fingerspelling is used extensively in some other sign languages such as ASL or BSL (Sutton-Spence 1994), while 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, generally it is quite simple for an avatar to fingerspell as fingerspelling includes no syntax, movement in signing space 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 fingerspelling or a code such as 'Signed Exact Greek' extensively. For these reasons the present project aims to represent a form of GSL as close to natural fluent signing as possible, and only uses fingerspelling occasionally, for example in language games, where teaching of written Greek is the focus.

4.2 Notation and glossing

In order to decide on the notation to be followed for sign recording in the lexical resources DB, the existing international systems of sign language recording were evaluated in respect to effectiveness as to determination of the intermediate language of the system (see also Pizzuto & Pietrandrea (2000), for a more theoretical discussion). The latter consists an important 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.

Tools for transcription and notation of GSL include HamNoSys, a pictographic notation system developed by the University of Hamburg for the description of the phonology of signs (Prillwitz et al., 1989). HML files in HamNoSys will form the corpus of GSL lemmas while for the representation of sequential structures (i.e. in the phrase level) ELAN language annotator developed by the Max-Planck Institute of Psycholinguistics in Nijmegen, the Netherlands, will be used. 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 morpho-phonological description, with one additional tier with written Greek words of harsh semantic equivalents of utterances. It is an aim of the project to add more tiers as the project continues, such as those mentioned above on the use of non-manual features and on pragmatics, using the esxisting symbols in HamNoSys and ELAN. Signwriting was another transcribing tool under consideration, but was not chosen, given the expected compatibility of HamNoSys within the Elan tiers in the near future.

5. Tutoring system description - corpus of educational material

The user interface under development is based on technologies (experience gained in previous SPERO and Faethon projects) which enable tracing the personal characteristics of specific users, on the basis of combination of personal data and his/her responses, previously acquired knowledge and user classification, so that the teaching process may be best customised. 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 will involve an initiator of the session, the student-s 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 whiteboard, icon banks and chat board visible in the screen along with the virtual signer for common use in the classroom. The participants will also be able 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 (Kourbetis & Efthimiou, 2003). 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 / written Greek. Concepts such as subject-object relations, types of verbs, discourse functions of the language form the units of the curriculum in SYNENNOESE so that the same principles are taught under the same umbrella, but without projecting onto GSL a mirror image of the Greek grammar. For the selection and arrangement of the educational material the project is in close cooperation with the Pedagogical Institute in Athens, which is the main official agency in charge of the development of educational material.

According to EU principles for accessibility to information in special education (see also WP COM (2000) 284 final), all Greek schools have been provided with suitable equipment for unrestricted Internet access, so the deliverables of the project can be readily applicable to real life school routine. Unfortunately, though, there have been no official educational resources for primary education of the deaf in the area of languages, until the time of writing of the current work. SYNENNOESE is the first applicable project for open and distance learning for the deaf, either individually or in group sessions. After month 12 of the beginning of the project there will be a trial period in sample student and tutor groups with the aid of the Pedagogical Institute for feedback and corrections.

6. Technical considerations

The implementation team has reviewed currently available avatar 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 friendly level of description, before they are transformed in parameters of body movement (Body Animation Parameters –BAPs) according to the MPEG-4 model. In the area of text-to-sign animation there have been some similar projects (VISICAST, Thetos, SignSynth and eSIGN among them) that SYNENNOESE uses as background.

Technologies considered for the viewing and interaction of 3D models wereVRML (Virtual Reality Modeling Language), X3D (eXtensible 3D) and H-ANIM. VRML (Virtual Reality Modelling Language) is a high level formal language with the ability to describe 3D interactive

objects and worlds. It is a hierarchical scene description language that defines the geometry and behaviour of a 3D scene or "world" and the way in which this is navigated by the user. VRML is the only standardised (ISO/IEC 14772) 3D format suitable for Web delivery.
X3D is the next-generation open standard for 3D on the web. It is an extensible standard that can easily be supported by content creation tools, proprietary browsers, and other 3D applications, both for importing and exporting. It replaces VRML, but also provides compatibility with existing VRML content and browsers.
H-ANIM is a set of specifications for description of human animation, based on body segments and connections. According to the H-ANIM standard, the human body consists of a number of segments (such as the forearm, hand and foot), which are connected to each other by joints (such as the elbow, wrist and ankle). H-ANIM can be used to describe the gestures. Motion tracking and haptic devices (such as CyberGrasp or Acceleration Sensing Glove with a virtual keyboard) were initially considered but it was agreed that, if quality of the results of the first transcribed signs with application of HamNoSys notation commands is acceptable, motion capture sequences will not need to be applied. In either case, both are much more flexible solutions than using 'frozen' mpeg or avi video files. Avatars are much more accessible to flexible information exchange and take advantage of the dynamic nature of phonological and syntactic rules.

7. Adopted 3D technologies

For the content designer to interact with an avatar, a scripting language is required. In our implementation, we chose the STEP language (Scripting Technology for Embodied Persona) (Huang, Eliens & Visser (2002)). as the intermediate level between the end user and the virtual actor. A major advantage of 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 avatar; as a result, one may alter the definition of any action, without the need to re-model the virtual actor. The avatars that are utilized here, are compliant with the H-ANIM standard, so one can use any of the readily available or model a new one.

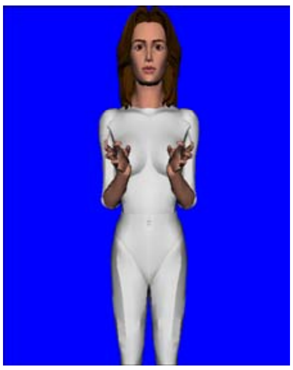


Figure 1: The virtual signer signing "radio" in GSL

An integrated system based on STEP is usually deployed in a usual HTML page, in order to maximize interoperability and be accessible to as many users as possible. This page includes an embedded VRML object, which represents the avatar and includes references to the STEP engine and the related JavaScript interface. From this setup, one may choose to create one's own script, for sign representation, and execute them independently, or embed them as JavaScript code, for maximized extensibility. The common VRML viewing plug-ins offer the possibility to select the required viewpoint at run-time, so it is possible for the user to experience the signing from any desired point of view (Kennaway, 2001 ; Kennaway, 2003 ; Huang, Eliens, & Visser, 2002). As an example, a frame of the signing sequence for "radio" is presented in figure 1.
In SYNENNOESE, a syntactic parser decodes the structural patterns of written Greek and matches them into the equivalents in GSL (Boutsis et al., 2000), and these resulting patterns are signed by a virtual human (avatar).
Using the technologies above, an internet platform will make access easy and fast, while the use of animated models instead of video files saves valuable storage space and bandwidth. Other advantages are the possibility of preview of predefined movements of the humanoid and the possibility of adding new movements and handshapes onto the system at any moment (script authoring). The advantages of an H-ANIM model (used version is v. 1.1) are its compatibility with VRML 97, flexibility on all segments and a more straightforward use.
The chart below (Figure 2) shows how the system functions and how data is transferred between machine and users. The testbed includes a page with and embedded VRML97 object, a JavaScript form for communication with the user and a Java Applet for communication with the back-end system. As can be seen in the chart, the system does not involve recognition of speech or signs. Machine translation mechanisms are at the background while at the present the output is a medium for human to non-human communication, rather than a machine for automatic translation.

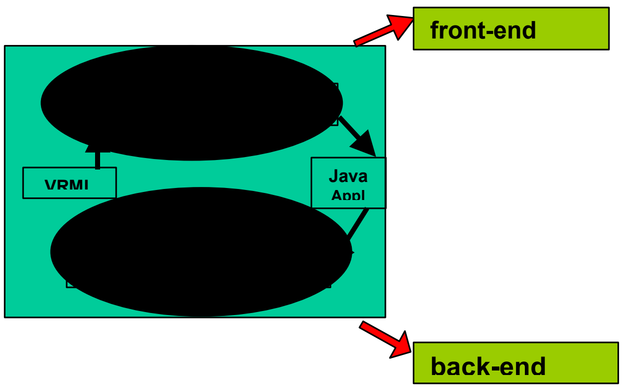


Figure 2. Data flow chart

8. Implications and extensibility of the educational platform

- As an educational tool above all, SYNENNOESE offers a user-friendly environment for young deaf pupils aged 6 to 9 so they can have visual translation of words and phrases. The signed feedback acts as a motivating tool for spelling Greek words and structuring sentences correctly, as well for evaluating one's performance. For deaf young

students as a group with special needs, the platforms 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 SYNENNOESE may receive unlimited educational content besides primary school grammar units. On the other hand, unlimited school units, such as the increasing special units with individual deaf students in rural areas and islands can link with one another via SYNENNOESE.
- 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.
- 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 the major outcome of the whole project. In this way the representation of GSL structures can be matched to equivalents ones of written Greek, and it will be a challenge to be able 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 for research across signed languages.
- From a socio-economic point of view, creating this platform will greatly contribute towards the inclusion of deaf people in Greek society in an environment of equal opportunities.

9. 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 projects, and they were expected before the beginning of the project.

From a linguistic and educational point of view, the major issues that need to be addressed are the following:

- 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 GSL dictionaries in DVD form with avi video entries. Moreover, the platform will be open to updates through the script authoring process.
- A second problem 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 above on grammar content definition, monomorphemic entries were agreed to be included in the first stage. In the next stages there will be 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.
- The data available in GSL, when compared with data from 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 Pedagogical Institute and the feedback from the Deaf Community, along with the continuing data from GSL linguistic research.
- Lastly, all schools in Greece have recently become accessible to the Internet, Deaf settings included. In practice however, there are many more accessibility barriers for a considerable number of deaf students who have additional special needs. Relevant provisions have been made according to general accessibility principles for these students (as to text size, keyboard settings etc) but the pilot application of the platform in December 2004 after 12 months of the beginning of the project will certainly indicate more points for development.

Technical problems include:

- A solution for smooth transition between signs and fusion between handshapes so that neighboring signs in a sentence appear as naturally articulated as possible.
- Automated commands for grammatical use of 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, as all of the above non manually signed features are considered as internalized parts of GSL grammar.
- It would be ideal to have a readily available system for retrieving and automatically extend phonological rules via HamNoSys notation. To the best of our knowledge such provisions are being made and the problem will meet a solution soon.
- 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 teams involved in the SYNENNOESE 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.

10. Conclusion

Given that the platform under discussion consists an original research object, successful completion of its development will open the way to a complete support system for the education of the Deaf Community members in Greece.

Acknowledgments

The authors wish to acknowledge the assistance of all groups involved in the implementation of the discussed platform: computational linguists, computer scientists, GSL specialists, the panel of GSL consultants and the representatives of the Greek Deaf Community, acting as our informants. This work was partially supported by the national project grant SYNENNOESE (GSRT: e-Learning 44).

References

Antzakas, K. & Woll, B. (2001). Head Movements and Negation in Greek Sign Language. Gesture Workshop, City University of London 2001: 193--196.

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

Boutsis S., Prokopidis P., Giouli V. & Piperidis S., (2000). A Robust Parser for Unrestricted Greek Text. In Proceedings of the 2nd Language Resources and Evaluation Conference, pages 467-474.

Boyes Braem, P. & Sutton-Spence, R. (eds.) (2001). The Hands are the Head of the Mouth: the Mouth as Articulator in Sign Languages Hamburg : Signum, c2001. vi, 291 p. International studies on sign language and communication of the deaf ; v.39.

Brennan, M. (1987). British Sign Language, the Language of the Deaf Community, in T. Booth & W. Swann (eds), Including People with Disabilities: Curricula for All, Open University Press: Milton Keynes.

Brien D. & Brennan, M. (1992). Dictionary of British Sign Language / English) Faber and Faber, London Boston.

Bybee, J. (1985). Morphology: a Study of the Relation between Meaning and Form. John Benjamins publishing, Amsterdam / Philadelphia.

Comrie, B. (1981). Language Universals and Linguistic Typology : Syntax and Morphology /Bernard Comrie. Oxford: Blackwell.

Efthimiou, E. & Katsoyannou, M. (2001). Research issues on GSL: a study of vocabulary and lexicon creation. Studies in Greek Linguistics, Vol. 2 Computational Linguistics, 42--50 (in Greek).

Efthimiou, E. & Katsoyannou, M. (2002). NOEMA: a Greek Sign Language Modern Greek bidirectional dictionary. Modern Education Vol. 126/127, 2002, 115--118 (in Greek).

Efthimiou, E., Sapountzaki, G., Carpouzis, C., Fotinea, S-E. (2004). Developing an e-Learning platform for the Greek Sign Language. Lecture Notes in Computer Science (LNCS), Springer-Verlag Berlin Heidelberg (in print).

Engberg – Pedersen, E. (1993). Space in Danish Sign Language, Signum Press, Hamburg.

Gee, J. & Goodhart, W. (1985). Nativization, Linguistic Theory, and Deaf Language Acquisition. Sign Language Studies, 49, 291--342.

Greenberg J. H. ed. (1968) Universals of Human Language, MIT Press.

Huang, Z., Eliens, A., & Visser, C. (2002). STEP: A Scripting Language for Embodied Agents, Proceedings of the Workshop on Lifelike Animated Agents.

Johnston T. & Schembri, A. (1999). On defining Lexeme in a Signed Language. Sign Language and Linguistics 2:2, 115--185.

Kennaway, R. (2001). Synthetic Animation of Deaf Signing Gestures. International Gesture Workshop, City University, London.

Kennaway, R. (2003). Experience with, and Requirements for, a Gesture Description Language for Synthetic Animation. 5th International Workshop on Gesture and Sign Language based Human-Computer Interaction, Genova.

Kourbetis, V. & Efthimiou, E. (2003). Multimedia Dictionaries of GSL as language and educational tools. Second Hellenic Conference on Education, Syros (in Greek), (in print).

Kourbetis, V. (1999). Noima stin Ekpaideusi Athens, Greece, Hellenic Pedagogical Institute.

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

Lampropoulou, V. (1992). Meeting the needs of deaf children in Greece. A systematic approach. Journal of the British Association of Teachers of the Deaf, 16, 33--34.

Lampropoulou, V. (1997). I Ereuna tis ellinikis Noimatikis Glossas: Paratiriseis Phonologikis Analisis. Athens, Greece: Glossa.

Liddell, S. (1980). American Sign Language Syntax, The Hague: Mouton.

Logiadis, N. & Logiadis, M. (1985). Dictionary of Sign Language. Potamitis Press (in Greek).

Neidle, C., Kegl, J. et al. (2000). The Syntax of ASL. Functional Categories and Hierarchical Structure MIT Press, Cambridge, Massachusetts: London.

Padden, C. (1989). The Relationship between Space and Grammar in American Sign Language Verb Morphology, in Theoretical Issues in Sign Language Research, Gallaudet Univ. Press.

Pizzuto, E. & Pietrandrea, P. (2000). Notating signed texts: open questions and indications for further research (unpublished manuscript).

Prillwitz et al. (1989). HamNoSys. Version 2.0. Hamburg Notation System for Sign Language. An Introductory Guide. Broschur / Paperback (ISBN 3-927731-01-3).

PROKLISI Project: WP9: Development of a GSL based educational tool for the education of people with hearing impairments, Deliverable II: GSL linguistic material: data collection methodology. ILSP, May 2003.

Stokoe W & Kuschel R (1978). For Sign Language Research, Linstock Press.

Sutton-Spence, R. (1994). The Role of the Manual Alphabet and Fingerspelling in British Sign Language. PhD Dissertation. University of Bristol.

Sutton-Spence, R. & Woll, B. (1999). The Linguistics of British Sign Language; an Introduction, Cambridge University Press.
Valli, C., and Lucas, C. (1995). Linguistics of American Sign Language, 2nd ed. Washington D.C.: Gallaudet University Press.
Wilcox, S., Scheibmann, J., Wood, D., Cokely, D. & Stokoe, W. (1994). Multimedia dictionary of American Sign Language. In Proceedings of ASSETS Conference, Association for Computing Machinery, (pp. 9--16).
http://www.leidenuniv.nl/hil/sign-lang/slsites.html#technical
http://www.sign-lang.uni-hamburg.de/Quellen/default.html
http://www.fhs-hagenberg.ac.at/mtd/projekte/FFF/3dSign/bookmarks.html

Sharing sign language corpora online: proposals for transcription and metadata categories

Onno Crasborn[*], Els van der Kooij[*], Daan Broeder[†], Hennie Brugman[†]

[*] Department of Linguistics, University of Nijmegen
PO Box 9103, NL-6500 HD Nijmegen, The Netherlands
{o.crasborn, e.van.der.kooij}@let.kun.nl
[†] Technical group, Max Planck Institute for Psycholinguistics
PO Box 310, 6500 AH Nijmegen, The Netherlands
{daan.broeder, hennie.brugman}@mpi.nl

Abstract

This paper presents the results of a European project called ECHO, which included an effort to publish sign language corpora online. The aim of the ECHO project was to explore the intricacies of sharing data using the internet in all areas of the humanities. For sign language, this involved adding a specific profile to the IMDI metadata set for characterizing spoken language corpora, and developing a set of transcription conventions that are useful for a broad audience of linguists. In addition to presenting these results, we outline some options for future technological developments, and bring forward some ethical problems relating to publishing video data on internet.

1. The ECHO project

Within the EU project 'European Cultural Heritage Online' (ECHO)[1], one of the five case studies is devoted to the field of language studies. The case study is titled 'Language as cultural heritage: a pilot project with sign languages'[2]. New data have been recorded from three sign languages of different Deaf communities in Europe: Sign Language of the Netherlands (abbreviated SLN), British Sign Language (BSL) and Swedish Sign Language (SSL). By having people retell written fable stories, comparable data resulted that can be used for cross-linguistic research. In addition to these semi-spontaneous data, we have elicited basic word lists and included some sign language poetry (some newly recorded, some already published).

The first aim of this paper is to characterize the conventions that were used and to explain why these can be considered as useful to a larger audience of linguists. The ELAN and IMDI software tools that were used to enter the transcriptions and metadata store their data in XML files whose format is described by open schemata and which can be accessed by other software tools as well. Using these open-standard tools, we developed a set of transcription conventions that are likely to be useable by a large group of researchers with diverse interests.

The second aim of this paper is to outline some desired functionalities of these tools that will make it more attractive to actually use existing corpora. Finally, we will outline some ethical challenges that have not yet received much discussion in the sign language field

2. The need for standardization

For actual cross-linguistic studies to take place, it is necessary that not only the same stimulus material is used, or otherwise comparable data are used, but also that the same conventions for annotating these data are used, both in terms of linguistic transcription and in terms of metadata description. The availability of a small corpus of video recordings from different languages, as published for the ECHO project, hopefully promotes standardization.

2.1 Metadata standards

In our case, metadata descriptions of language corpora characterize the documents and data files that make up the corpus in terms of descriptors that pertain to the whole unit of media and transcription files, rather than to individual sections within the files. For example, information about the subjects, the identity of the researchers involved in the collection and the register used by the speakers or signers typically belongs to the metadata domain. Users can then search within and across large corpora for all transcribed video material with male signers older than 45 years, for example. However, for such searches to be possible, it is essential that users obey the same conventions for labeling corpora. A proposal for such a standard is presented in section 3[3]. This is a specialization of the IMDI set of metadata descriptors for language resources[4].

2.2 Transcription standards

Several tools are currently available for annotating video data. Both SyncWriter (Hanke & Prillwitz 1995) and SignStream[5] have developed especially for sign language data, whereas ELAN started its life in the domain of gesture research (former versions were called MediaTagger)[6].

These new technologies for presenting sign language data and transcriptions pose the following question: to what extent should we use standard transcription conventions? If all the raw material (the video sources)

[1] http://echo.mpiwg-berlin.mpg.de/
[2] http://www.let.kun.nl/sign-lang/echo/; project partners were the University of Nijmegen, City University London, and Stockholm University.
[3] Further information on the proposed standard can be found at http://www.let.kun.nl/sign-lang/IMDI/.
[4] http://www.mpi.nl/IMDI
[5] http://www.bu.edu/asllrp/SignStream/
[6] http://www.mpi.nl/tools/elan.html

is available, do we need full transcriptions? In principle, one can look at the video source for all kinds of information that are traditionally included in various transcription system, such as eye gaze, head nods, etc. On the other hand, the great strength of computer tools such as ELAN is that it allows for complex searches in large data domains and for the immediate inspection of the video fragments relating to the search results; this is typically very time consuming when using paper transcription forms or even digitized transcription forms that are not directly linked to the original data.

Within the ECHO project, we therefore wanted to establish an annotation system that could be useful for any researcher, with a focus on the syntactic and discourse domains. We tried to be careful not to impose too much analysis on any tier by saying that a specific phonetic form is an instance of 'person agreement', for example. On the other hand, analytical decisions are constantly being made in any transcription process. For example, even adding multiple tiers with translations in various written languages (in the case of the ECHO project: Dutch, English and Swedish) implies taking (implicit or explicit) decisions about where sentence boundaries are located.

While every research project will have its own research questions and require special transcription categories, it should be possible to define a standard set of transcription tiers and values that are useful to large groups of researchers, regardless of their specific interests. For example, a translation at sentence level to a written language is always useful, if only for exploring a video recording. Working with three teams of linguists from different countries, each with their own research interests, the ECHO project formed a good start for developing such a standard set of transcription conventions. This ECHO set is described in section 4.

The relatively small set of transcription tiers allows for the coding of a relatively large data set, which can be further expanded by researchers according to their specific needs. ELAN will see several updates in the near future; one of the future functions will be the possibility to expand a publicly available transcription file with ones own additions, including extra tiers, and storing these additions in a local file while maintaining the link to the original transcription that will be stored on a remote server.

3. Metadata description of sign language corpora: expanding the IMDI standard

3.1 The IMDI standard and profiles

The set of IMDI metadata descriptors that was developed for spoken language corpora distinguishes 7 categories for each session unit:

1. *Session*. The session concept bundles all information about the circumstances and conditions of the linguistic event, groups the resources (for example, video files and annotation files) belonging to this event, and records any administrative information for the event.
2. *Project*. Information about the project for which the sessions were originally created.
3. *Collector*. Name and contact information for the person who collected the session.
4. *Content*. A set of categories describing the intellectual content of the session.
5. *Actors*. Names, roles and further information about the people involved in the session, including the signers and addressees, but also, for example, the researchers who collected and transcribed the material.
6. *Resources*. Information about the media files, such as URL, size, etc.
7. *References*. Citations and URLs to relevant publications and other archive resources.

Each of these seven categories allow for extension by users, in the form of 'key–value pairs'. A key specifies an extra category, an extra field, for which a value can be specified. For example, one might specify a key called *Backup Copy* to quickly specify whether a back-up copy of the original tape has already been made (yes vs. no).

In a workshop for the ECHO project, held at the University of Nijmegen in May 2003, a group of sign linguists from various countries and with varying research interests sat together to see how these categories could be applied to sign language data. The outcome of that workshop was a set of key fields to describe sign language corpora. These extra categories have now been bundled in an extension to the standard IMDI metadata specification, called 'sign language profile'. Profiles in the IMDI Editor tool offer sets of extra fields that apply to specific types of data, in this case communication in a specific modality.

3.2 The sign language profile

The sign language profile adds key fields in two areas in the IMDI set: content and actors. All of the fields can be specified or left empty.

In content, *Language Variety* describes the specific form of communication used in the session, and *Elicitation Method* specifies the specific prompt used to elicit the data at hand. A set of four keys describes the communication situation with respect to interpreting: who was the interpreter (if any) interpreting for (*Interpreting.Audience*), what were the source and target modalities (*Interpreting.Source* and *Interpreting. Target*), and is the interpreter visible in the video recording (*Interpreting.Visibility*)?

Secondly, four sets of keys are defined that can be used to describe various properties of each actor who is related to the session: properties pertaining to deafness, the amount of sign language experience, the family members, and the (deaf) education of the actor.

Deafness.Status describes the hearing status of the actor (deaf, hard-of-hearing, hearing), and *Deafness.AidType* describes the kind of hearing aid the actor is using (if any).

The amount of *Sign Language Experience* is expressed by specifying the *Exposure Age*, *Acquisition Location* and experience with *Sign Teaching*.

The family of the actor can be described by specifying *Deafness* and *Primary Communication Form* for *Mother*, *Father* and *Partner*.

Finally, the *Education* history of the actor can be specified in a series of keys: *Age* (the start and end age

of the actor during his education), the *School Type* (primary school, university, etc.), the *Class Kind* (deaf, hearing, etc.), the *Education Model*, the *Location*, and whether the school was a *Boarding School* or not.
A more complete definition of the whole sign language profile is given in Crasborn & Hanke (2003).

4. A standard set of linguistic transcription conventions for sign language data

4.1 An introduction to ELAN and the 'tier' concept

Below we describe the different tiers used for the ECHO project[7]. A tier is a set of annotations that share the same characteristics, e.g. one tier containing all the glosses for the right hand and another tier containing the Dutch translations. ELAN distinguishes between two types of tiers: "parent tiers" and "child tiers". Parent tiers are independent tiers, which contain annotations that are linked directly to a time interval of the media frame. Child tiers or referring tiers contain annotations that are linked to annotations on another tier (the parent tier)[8]. ELAN provides the opportunity to select one or more video frames and assign a specific value to this selected time span. For example, when the eye brows are first up and then down (neutral) in the same sign, one would only select the time interval in the video in which the eyebrows are up for the brows tier, and mark that time-domain with a specific code (for instance 'up'). This is possible for all tiers that one creates.
It is important to emphasize that, unlike in the IMDI software, there is no standard set of tiers for any document. Tiers have to be set up by the user for every annotation file that is created to annotate a media file. The set that we propose is just that: a proposal for a set of tiers that cover elementary transcription categories that can be useful for many different kinds of research. The use of this set of tiers is exemplified by the data transcribed for the ECHO project[9]. Any user can add both additional tiers and additional annotations on existing tiers to the documents that have been published in the context of the ECHO project.

4.2 Tiers with general information

General information that can be supplied for every fragment of a video file includes *Translation* tiers for English, Swedish and Dutch. Each of these tiers target a translation at sentence level. An annotation on the *Role* tier indicates that the signer takes on the role of a specific discourse participant, as commonly happens in sign language discourse. Finally, the *Comments/notes* tier can be used to add any kind of comment by the user.

4.3 Tiers with manual information

Manual behavior is systematically described separately for the two hands. For both the left and the right hand, there is a *Gloss* tier. Child tiers for each of these two articulators specify whether there is *Repetition* in the movement of the glossed unit, and what the *Direction & Location* of each of the hands is.

4.4 Tiers with non-manual information

A set of non-manual tiers allow for the specification of some of the relevant properties of the face, head, and body of the signer. Movement of the *Head* and *Eye Brows* can be specified, as well as the amount of *Eye Aperture* (including the notation of eye blinks) and the direction of *Eye Gaze*.
A new system was devised to specify the behavior of the *Mouth*, including the tongue, which in previous systems was often treated in a rather fragmentary manner (Nonhebel, Crasborn & van der Kooij 2004b).

4.4 Properties of the transcription conventions

The transcription system outlined in the sections above had two central goals. First of all, it should be easy and relatively quick to use for encoders, so that users can transcribe considerable amounts of data within a reasonable time frame. This inevitably goes at the expense of detail. For example, for facial expression, the FACS system (Ekman, Friesen & Hager 2002) is the most detailed and accurate transcription method that is known, but it is extremely time-intensive to learn to master and use, and offers far more detail than is necessary for the large majority of research projects. The tiers for non-manual activity that we propose aim to form an optimal compromise between the amount of detail available to the user and the time investment made by the transcriber.
Secondly, we tried to systematically separate form from function for all tiers. Since the function of a given linguistic form can vary from language to language, it is crucial to emphasize the coding of the form of linguistic behavior.

5. Specifications for future tools

Most importantly in the context of this paper, searching across both data and metadata domains will need to be an important target of further development. In the present state of the tools, one needs to first search within the set of metadata categories, and in the resulting set of transcription files search for data categories one-by-one. Finding all cases of weak hand spreading by people younger than 20 thus becomes a very time-consuming task, whereas corpora are particularly useful for those kinds of complex queries.
In the sign language research community, working with corpus data is still very uncommon, presumably in part because there are no commonly used written forms of sign languages until now that have allowed to create text corpora. Now the computer technology is available to build up corpora of digitized video recordings and annotate these, in addition to the search facilities, software is needed to provide basic statistical functions in ELAN, including frequencies of annotation values on

[7] An extensive description is available in Nonhebel, Crasborn & van der Kooij (2004a).
[8] See also ELAN manual, available at http://www.mpi.nl/tools/elan.html.
[9] These data can be freely downloaded from http://www.let.kun.nl/sign-lang/echo/data.html.

different tiers and the distribution of the durations of these annotation values. Currently, the most obvious way to perform quantitative analyses of transcription files at this moment is to export data to a spreadsheet program for further analysis.

A function that is currently being implemented is to add a visualization of kinematic recordings with the transcription of video material, similar to the display of the oscillogram of sound files in ELAN. These numerical data can then be more easily integrated with qualitative analyses based on transcription. Additionally, the software will need to provide numerical analyses appropriate to phonetic analysis of sign languages, similar to the 'Praat' software for speech analysis (Boersma & Weenink 2004). As the field of sign language phonetics is still in its infancy, the specifications of such functionality will have to develop over the years to come. Finally, a similar integration of quantitative data from eye-tracking equipment would enhance the usability of the software for some research groups.

Working together with colleagues anywhere in the world on the same annotation document at the same time is another function currently under development. Using peer-to-peer technology, it will become possible to look at the same annotation document on different computers connected to the internet, and instantly see modifications that are being made by the other party. In combination with a chat function, one can jointly look at existing annotations and create new annotations (see Brugman, Crasborn & Russel 2004 for further details on this 'collaborative annotation' concept).

6. Ethical aspects of publishing sign language data online

Needless to say, the privacy of subjects in scientific studies has to be respected. For the sign language study in the ECHO project, this gives rise to extra problems not previously encountered in the creation of spoken language corpora that just make use of sound recordings. The visual information in the video recordings contains a lot more personal information than audio recordings of voices, including not only the identity of the signer (i.e., the visual appearance of the face), but also more clues to the emotional state and age of the person, for example.

While it is common practice to ask subjects in linguistic recordings for their explicit written permission to use the recordings for various purposes, including making images for publications, discussion among sign language specialists revealed that this permission is a rather sensitive issue in the case of internet publication. Publication of data online imply that the information is available to the whole world, and not just to a limited group of people with access to specific university libraries, for example, as in the case of video tape recordings used until recently. Signers who have no problem with the inclusion of the video data at the time of recording may well regret this choice 15 years later. Can this be considered the problem of the person involved, or should researchers make more of an effort to outline the implications of sharing data to subjects? Alternatively, data access can be restricted to linguists registered as users of the corpus by the host institution, but this comes down to restricting access to data that were intended to be public – at least within the open access concept that is central to the ECHO project.

Future projects aimed at making data accessible online should explore these issues in more depth, with assistance from both legal and ethics specialists.

7. References

Boersma, P. & D. Weenink (2004) Praat. Doing phonetics by computer. http://www.praat.org.

Brugman, H., O. Crasborn & A. Russel (2004) Collaborative annotation of sign language data with peer-to-peer technology. Paper presented at LREC 2004, Lissabon.

Crasborn, O. & T. Hanke (2003) Additions to the IMDI metadata set for sign language corpora. http://www.let.kun.nl/sign-lang/echo/docs/SignMetadata_Oct2003.doc.

Ekman, P., W. Friesen & J. Hager (2002) *Facial Action Coding System*. Salt Lake City: Research Nexus.

Hanke, T. & S. Prillwitz (1995) syncWRITER: Integrating video into the transcription and analysis of sign language. In H. Bos & T. Schermer (eds.) *Sign language research 1994*. Hamburg: Signum. Pp. 303-312.

Nonhebel, A., O. Crasborn & E. van der Kooij (2004a) Sign language transcription conventions for the ECHO project. http://www.let.kun.nl/sign-lang/echo/docs/transcr_conv.pdf.

Nonhebel, A., O. Crasborn & E. van der Kooij (2004b) Sign language transcription conventions for the ECHO project. BSL and NGT mouth annotations. http://www.let.kun.nl/sign-lang/echo/docs/transcr_mouth.pdf.

Spatial Representation of Classifier Predicates for Machine Translation into American Sign Language

Matt Huenerfauth

Computer and Information Science
University of Pennsylvania
Philadelphia, PA 19104
matthewh@seas.upenn.edu

Abstract

The translation of English text into American Sign Language (ASL) animation tests the limits of traditional machine translation (MT) approaches. The generation of spatially complex ASL phenomena called "classifier predicates" motivates a new representation for ASL based on virtual reality modeling software, and previous linguistic research provides constraints on the design of an English-to-Classifier-Predicate translation process operating on this representation. This translation design can be incorporated into a multi-pathway architecture to build English-to-ASL MT systems capable of producing classifier predicates.

Introduction and Motivations

Although Deaf students in the U.S. and Canada are taught written English, the challenge of acquiring a spoken language for students with hearing impairments results in the majority of Deaf U.S. high school graduates reading at a fourth-grade[1] level (Holt, 1991). Unfortunately, many strategies for making elements of the hearing world accessible to the Deaf (e.g. television closed captioning or teletype telephone services) assume that the user has strong English literacy skills. Since many Deaf people who have difficulty reading English possess stronger fluency in American Sign Language (ASL), an automated English-to-ASL machine translation (MT) system can make more information and services accessible in situations where English captioning text is at too high a reading level or a live interpreter is unavailable.

Previous English-to-ASL MT systems have used 3D graphics software to animate a virtual human character to perform ASL output. Generally, a script written in a basic animation instruction set controls the character's movement; so, MT systems must translate English text into a script directing the character to perform ASL. Previous projects have either used word-to-sign dictionaries to produce English-like manual signing output, or they have incorporated analysis grammar and transfer rules to produce ASL output (Huenerfauth, 2003; Sáfár and Marshall, 2001; Speers, 2001; Zhao et al., 2000). While most of this ASL MT work is still preliminary, there is promise that an MT system will one day be able to translate many kinds of English-to-ASL sentences; although, some particular ASL phenomena – those involving complex use of the signing space – have proven difficult for traditional MT approaches. This paper will present a design for generating these expressions.

ASL Spatial Phenomena

ASL signers use the space around them for several grammatical, discourse, and descriptive purposes. During a conversation, an entity under discussion (whether concrete or abstract) can be "positioned" at a point in the signing space. Subsequent pronominal reference to this entity can be made by pointing to this location (Neidle et al., 2000). Some verb signs will move toward or away from these points to indicate (or show agreement with) their arguments (Liddell, 2003a; Neidle et al., 2000). Generally, the locations chosen for this use of the signing space are not topologically meaningful; that is, one imaginary entity being positioned to the left of another in the signing space doesn't necessarily indicate the entity is to the left of the other in the real world.

Other ASL expressions are more complex in their use of space and position invisible objects around the signer to topologically indicate the arrangement of entities in a 3D scene being discussed. Constructions called "classifier predicates" allow signers to use their hands to position, move, trace, or re-orient an imaginary object in the space in front of them to indicate the location, movement, shape, contour, physical dimension, or some other property of a corresponding real world entity under discussion. Classifier predicates consist of a semantically meaningful handshape and a 3D hand movement path. A handshape is chosen from a closed set based on characteristics of the entity described (whether it is a vehicle, human, animal, etc.) and what aspect of the entity the signer is describing (surface, position, motion, etc).

For example, the sentence "the car drove down the bumpy road past the cat" could be expressed in ASL using two classifier predicates. First, a signer would move a hand in a "bent V" handshape (index and middle fingers extended and bent slightly) forward and slightly downward to a point in space in front of his or her torso where an imaginary miniature cat could be envisioned. Next, a hand in a "3" handshape (thumb, index, middle fingers extended with the thumb pointing upwards) could trace a path in space past the "cat" in an up-and-down fashion as if it were a car bouncing along a bumpy road. Generally, "bent V" handshapes are used for animals, and "3" handshapes, for vehicles.

Generating Classifier Predicates

As the "bumpy road" example suggests, translation involving classifier predicates is more complex than most English-to-ASL MT because of the highly productive and spatially representational nature of these signs. Previous ASL MT systems have dealt with this problem by omitting these expressions from their

[1] Students who are age eighteen and older are reading English text at a level more typical of a ten-year-old student.

linguistic coverage; however, many English concepts lack a fluent ASL translation without them. Further, these predicates are common in ASL; in many genres, signers produce a classifier predicate on average once per 100 signs (this is approximately once per minute at typical signing rates) (Morford and MacFarlane, 2003). So, systems that cannot produce classifier predicates can only produce ASL of limited fluency and are not a viable long-term solution to the English-to-ASL MT problem.

Classifier predicates challenge traditional definitions of what constitutes linguistic expression, and they oftentimes incorporate spatial metaphor and scene-visualization to such a degree that there is debate as to whether they are paralinguistic spatial gestures, non-spatial polymorphemic constructions, or compositional yet spatially-parameterized expressions (Liddell, 2003b). No matter their true nature, an ASL MT system must somehow generate classifier predicates. While MT designs are not required to follow linguistic models of human language production in order to be successful, it is worthwhile to consider linguistic models that account well for the ASL classifier predicate data but minimize the computational or representational overhead required to implement them.

Design Focus and Assumptions

This paper will focus on the generation of classifier predicates of movement and location (Supalla, 1982; Liddell, 2003a). Most of the discussion will be about generating individual classifier predicates; an approach for generating multiple interrelated predicates will be proposed toward the end of the paper.

This paper will assume that English input sentences that should be translated into ASL classifier predicates can be identified. Some of the MT designs proposed below will be specialized for the task of generating these phenomena. Since a complete MT system for English-to-ASL would need to generate more than just classifier predicates, the designs discussed below would need to be embedded within an MT system that had other processing pathways for handling non-spatial English input sentences. The design of such multi-pathway MT architectures is another focus of this research project (Huenerfauth, 2004).

These other pathways could handle most inputs by employing traditional MT technologies (like the ASL MT systems mentioned above). A sentence could be "identified" (or intercepted) for special processing in the classifier predicate pathway if it fell within the pathway's implemented lexical (and – for some designs – spatial) resources.[2] In this way, a classifier predicate generation component could actually be built on top of an existing ASL MT system that didn't currently support classifier predicate expressions.

We will first consider a classifier predicate MT approach requiring little linguistic processing or novel ASL representations, namely a fully lexicalized approach. As engineering limitations are identified or additional linguistic analyses are considered, the design will be modified, and progressively more sophisticated representations and processing architectures will emerge.

Design 1: Lexicalize the Movement Paths

The task of selecting the appropriate handshape for a classifier predicate, while non-trivial, seems approachable with a lexicalized design. For example, by storing semantic features (e.g. +human, +vehicle, +animal, +flat-surface) in the English lexicon, possible handshapes can be identified for entities referred to by particular English nouns. Associating other features (e.g. +motion-path, +stationary-location, +relative-locations, +shape-contour) with particular verbs or prepositions in the English lexicon could help identify what kind of information the predicate must express – further narrowing the set of possible classifier handshapes. To produce the 3D movement portion of the predicate using this lexicalized approach, we could store a set of 3D coordinates in the English lexicon for each word or phrase (piece of lexicalized syntactic structure) that may be translated as a classifier predicate.

Problems with This Design

Unfortunately, the highly productive and scene-specific nature of these signs makes them potentially infinite in number. For example, while it may seem possible to simply store a 3D path with the English phrase "driving up a hill," factors like the curve of the road, steepness of hill, how far up to drive, etc. would affect the final output. So, a naïve lexicalized 3D-semantics treatment of classifier movement would not be scalable.

Design 2: Compose the Movement Paths

Since the system may need to produce innumerable possible classifier predicates, we can't merely treat the movement path as an unanalyzable whole. A more practical design would compose a 3D path based on some finite set of features or semantic elements from the English source text. This approach would need a library of basic animation components that could be combined to produce a single classifier predicate movement. Such an "animation lexicon" would contain common positions in space, relative orientations of objects in space (for concepts like above, below, across from), common motion paths, or common contours for such paths. Finally, these components would be associated with corresponding features or semantic elements of English so that the appropriate animation components can be selected and combined at translation time to produce a 3D path.

Problems with This Design

This design is analogous to the polymorphemic model of classifier predicate generation (Supalla 1978, 1982, 1986). This model describes ASL classifier predicates as categorical, and it characterizes their generation as a process of combining sets of spatially semantic morphemes. The difficulty is that every piece of spatial information we might express with a classifier predicate must be encoded as a morpheme. These phenomena can convey such a wide variety of spatial

[2] A later section of this paper describes how the decision of whether an input English sentence can be processed by the special classifier predicate translation pathway depends on whether a *motif* (introduced in that section) has been implemented for the semantic domain of that sentence.

information – especially when used in combination to describe spatial relationships or comparisons between objects in a scene – that many morphemes are required.

Liddell's analysis (2003b) of the polymorphemic model indicates that in order to generate the variety of classifier predicates seen in ASL data, the model would need a tremendously large (and possibly infinite) number of morphemes. Using a polymorphemic analysis, Liddell (2003b) decomposes a classifier predicate of one person walking up to another, and he finds over 28 morphemes, including some for: two entities facing each other, being on the same horizontal plane, being vertically oriented, being freely moving, being a particular distance apart, moving on a straight path, etc.

Liddell considers classifier predicates as being continuous and somewhat gestural in nature (2003a), and this partially explains his rejection of the model. (If there are not a finite number of possible sizes, locations, and relative orientations for objects in the scene, then the number of morphemes needed becomes infinite.) Whether classifier predicates are continuous or categorical and whether this number of morphemes is infinite or finite, the number would likely be intractably large for an MT system to process. We will see that the final classifier predicate generation design proposed in this paper will use a non-categorical approach for selecting its 3D hand locations and movements. This should not be taken as a linguistic claim about human ASL signers (who may indeed use the large numbers of morphemes required by the polymorphemic model) but rather as a tractable engineering solution to the highly productive nature of classifier predicates.

Another reason why a polymorphemic approach to classifier predicate generation would be difficult to implement in a computational system is that the complex spatial interactions and constraints of a 3D scene would be difficult to encode in a set of compositional rules. For example, consider the two classifier predicates in the "the car drove down the bumpy road past the cat" example. To produce these predicates, the signer must know how the scene is arranged including the locations of the cat, the road, and the car. A path for the car must be chosen with beginning/ending positions, and the hand must be articulated to indicate the contour of the path (e.g. bumpy, hilly, twisty). The proximity of the road to the cat, the plane of the ground, and the curve of the road must be selected. Other properties of the objects must be known: (1) cats generally sit on the ground and (2) cars generally travel along the ground on roads. The successful translation of the English sentence into these two classifier predicates involved a great deal of semantic understanding, spatial knowledge, and reasoning.

A 3D Spatial Representation for ASL MT

ASL signers using classifier predicates handle these complexities using their own spatial knowledge and reasoning and by visualizing the elements of the scene. An MT system may also benefit from a 3D representation of the scene from which it could calculate the movement paths of classifier predicates. While design 2 needed compositional rules (and associated morphemes) to cover every possible combination of object positions and spatial implications as suggested by English texts, the third and final MT design (discussed in a later section) will use virtual reality 3D scene modeling software to simulate the movement and location of entities described by an English text (and to automatically manage their interactions).

The AnimNL System

A system for producing a changing 3D virtual reality representation of a scene from an English text has already been implemented: the Natural Language Instructions for Dynamically Altering Agent Behaviors system (Schuler, 2003; Bindiganavale et al., 2000; Badler et al., 2000) (herein, "AnimNL"). The system displays a 3D animation and accepts English input text containing instructions for the characters and objects in the scene to follow. It updates the virtual reality so that objects obey the English commands. AnimNL has been used in military training and equipment repair domains and can be extended by augmenting its library of Parameterized Action Representations (PARs), to cover additional domains of English input texts.

The system's ability to interact with language and plan future actions arises from the use of PARs, which can be thought of as animation/linguistic primitives for structuring the movements in a 3D scene. PARs are feature-value structures that have slots specifying: what agent is moving, the path/manner of this motion, whether it is translational/rotational motion, the terminating conditions on the motion, any speed or timing data, etc. A single locomotion event may contain several sub-movements or sub-events, and for this reason, PARs may be defined in a hierarchical manner. A single "high-level" PAR may specify the details for the entire motion, but it may be defined in terms of several "low-level" PARs which specify the more primitive sub-movements/events.

The system stores a database of PAR templates that represent prototypical actions the agent can perform. These templates are missing particular details (some of their slots aren't filled in) about the position of the agent or other entities in the environment that would affect how the animation action should really be performed in particular situations. By parameterizing PARs on the 3D coordinates of the objects participating in the movement, the system can produce animations specific to particular scene configurations and reuse common animation code.

English lexicalized syntactic structures are associated with PARs so that the analysis of a text is used to select a PAR template and fill some of its slots. For example, there may be a PAR associated with the concept of "falling" vs. another for "jumping." While these templates must remain parameterized on the 3D location of the agent of the movement until it is known at run time, there are some properties (in this case, the direction of motion) that can be specified for each from the English semantics. During analysis of the English input text, semantic features of motion verbs are obtained from the VerbNet hierarchy (Kipper et al., 2004), and these features are also used to select and fill a particular motion template. Since VerbNet groups verbs that share common semantic/syntactic properties, AnimNL is able to link an entire set of semantically similar motion verbs to a single PAR template. Each of the verbs in the set may fill some of the slots of the motion template somewhat differently.

When a PAR template has been partially filled with information from the English text and 3D object locations, it is passed off to AnimNL's animation planner. In fact, PARs contain slots allowing them to be hierarchical planning operators: pre-conditions, effects, subplans, etc. The movements of all objects in the AnimNL system are governed by a planning process, which allows the objects in the scene to move realistically. Many spatial motions have conditions on the location, orientation, or motion state of an object and its environment before, during, and after the event. The PAR operators help the system work out the details of an animation from the limited specification of this motion provided by an English text. For example, it may determine starting and stopping locations for movement paths or select relative locations for objects in the 3D scene based on prepositions and adverbials in the English input text. The interaction and conditions of these planning operators simulate physical constraints, collision avoidance, human anatomical limitations, and other factors to produce an animation.

Using AnimNL for ASL

The MT system's classifier predicate generator can use the AnimNL software to analyze English sentences to be translated into classifier predicates. AnimNL can process this text as if it were commands for the entities mentioned in the text to follow. Based on this analysis, the AnimNL can create and maintain a 3D representation of the location and motion of these entities. Next, a miniature virtual reality animation of the objects in this representation can be overlaid on a volume of the space in front of the torso of the animated ASL-signing character. In this way, a miniature 3D virtual reality would be embedded within the original 3D space containing the standing animated virtual human. In the "bumpy road" example, a small invisible object would be positioned in space in front of the chest of the signing character to represent the cat. Next, a 3D animation path and location for the car (relative to the cat) would be chosen in front of the character's chest.

The AnimNL software can thus produce a miniature "invisible world" representing the scene described by the input text. Unlike other applications of AnimNL – where entities described by the English text would need to be rendered to the screen – in this situation, the 3D objects would be transparent. Therefore, the MT system does not care about the exact appearance of the objects being modeled. Only the location, orientation, and motion paths of these objects in some generic 3D space are important since this information will be used to produce classifier predicates for the animated ASL-signing character.

An Overly Simplistic Generation Strategy

The next section of this paper (design 3) will discuss how the "invisible world" representation can be used to generate classifier predicates. To motivate that third and final design, we will first consider an overly simplistic (and incorrect) strategy for using the virtual reality to attempt classifier predicate generation.

This simplistic "Directly Pictorial" strategy for building a classifier predicate is as follows: When a new object is introduced into the invisible world, the signing character moves its hand to a location "inside of" the transparent object. By also choosing an appropriate handshape for the character (possibly using the +animal or +vehicle features discussed above), then a classifier predicate is apparently produced that conveys the spatial information from the English text. As objects in the invisible world are moved or reoriented as AnimNL analyzes more text, the signer can express this information using additional classifier predicates by again placing its hand inside the (possibly moving) 3D object. (See Figure 1.)

Limitations of the "Directly Pictorial" Strategy

Whereas design 2 mirrored the polymorphemic model, this design is similar to that of DeMatteo (1977), who sees classifier predicates as being direct "spatial analogues" of 3D movement paths in a scene imagined by the signer (Liddell, 2003b). In this model, signers maintain a 3D mental image of a scene to be described, select appropriate handshapes to refer to entities in their model, and trace out topologically analogous location and movement paths for these entities using their hands.

Unfortunately, the model is over-generative (Liddell, 2003b). By assuming that the selection of handshapes and movements are orthogonal and that movement paths are directly representative [3] of the paths of entities in space, this analysis predicts many ASL classifier constructions that never appear in the data (containing imaginable but ungrammatical combinations of handshape, orientation, and movement) (Liddell, 2003b). Finally, the model cannot consider discourse and non-spatial semantic features that can influence classifier predicate production in ASL.

Design 3: Lexicon of Classifier Predicates

The "Directly Pictorial" strategy was just one way to use the 3D information in the invisible world representation to generate classifier predicates. This section will introduce the MT approach advocated by this paper: design 3. This design uses the invisible world but avoids the limitations of the previous strategy by considering additional sources of information during translation. Whereas previous sections of this paper have used comparisons to linguistic models to critique an MT design, this section will use a linguistic model for inspiration.

Lexicon of Classifier Predicate Templates

Liddell (2003a, 2003b) proposed that ASL classifier predicates are stored as large numbers of abstract templates in a lexicon. They are "abstract" in the sense that each is a template parameterized on 3D coordinates of whatever object is being described, and each can therefore be instantiated into many possible

[3] To illustrate how classifier predicate movements can be conventional and not visually representative, Liddell (2003b) uses the example of an upright figure walking leisurely being expressed as a classifier predicate with D handshape slightly bouncing as it moves along a path. While the hand bounces, the meaning is not that a human is bouncing but that he or she is walking leisurely.

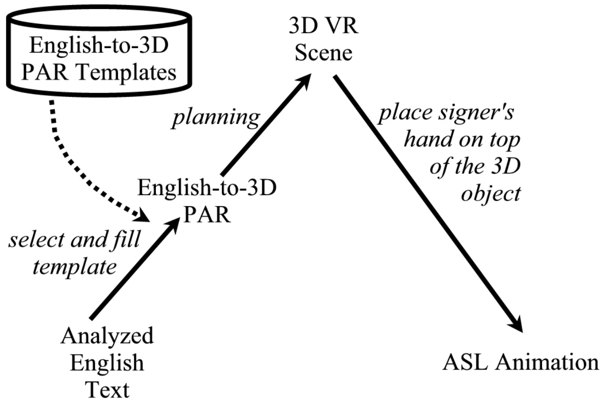


Figure 1: "Directly Pictorial" Generation Strategy (argued against in this paper). Solid lines depict transformation processes between representations, and dotted lines, information flow into a process.

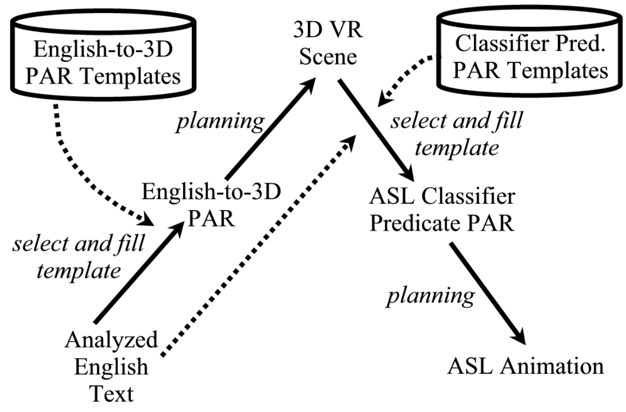


Figure 2: The Design 3 Architecture. Notice the new selection/filling process for a Classifier Predicate PAR based on: a PAR template, the 3D scene data, and English text features.

classifier predicate outputs. For example, there may be one template for classifier predicates expressing that a car is parked at a point in space; when this template is turned into an actual classifier predicate, then the 3D coordinate of the car would be filled in.

Each lexical entry stores the semantic content of a particular classifier predicate and most of the handshape and movement specification for its performance. A signer selects a template based on how well its spatial and non-spatial semantics convey the desired content. When a signer generates a classifier predicate from this template, then the locations, orientations, and specific movement paths of objects in a 3D mental spatial representation are used to fill the remaining parameters of the template and produce a full specification of how to perform the classifier predicate.

Although the previous paragraph refers to this approach as "lexical," it differs from design 1 (which augmented the English lexicon with 3D movement data) because it creates a distinct ASL lexicon of classifier predicates, and the movement information in these entries is parameterized on the data in the 3D scene. While these templates may also resemble the compositional morphemes of the polymorphemic model (the "animation lexicon" of design 2) since they both link semantics to 3D movement, these templates have more pre-compiled structure. While the morphemes required complex processing by compositional rules, the templates just need to be selected and to have their 3D parameters set.

Liddell (2003b) explains that this model avoids the under-generation of (Supalla, 1978, 1982, 1986) by incorporating a 3D spatial representation to select locations and movement paths, but it also avoids the over-generation of (DeMatteo, 1977) by restricting the possible combinations of handshapes and movement paths. Impossible combinations are explained as lexical gaps; ungrammatical classifier predicate feature combinations are simply not entries in the lexicon (Liddell, 2003b).

Classifier Predicate Templates for MT

To implement this linguistic model as an MT design, we will need: (1) a 3D scene representation, (2) a list of templates for producing the signing character's arm movements, (3) a way to link the semantics of English sentences to specific templates, and (4) a method for turning a filled template into an animation of the signer's arm. Requirement 1 is satisfied by the invisible world representation produced by the AnimNL software.

While the AnimNL software used one database of PAR templates to produce the 3D animation of objects in the invisible world, this design can fulfill requirement 2 by adding a second database, whose PAR templates will describe the animated movement of the signing character's arm as it performs a classifier predicate. (This first set will be called "invisible world" PARs, and the second, "classifier predicate" PARs.) Compared to the invisible world PARs, the classifier predicate PARs will be very simple: they will store instructions for the signing character's hand to be in a particular shape and for it move between two or more 3D coordinates in the signing space – possibly along a programmed contour.

The re-use of PAR templates suggests a method for linking the semantics of the English text to arm movement templates (requirement 3). Just as the AnimNL software used features of lexical syntactic structures to trigger invisible world PARs, design 3 can use these features to link the semantics of English sentences to classifier predicate PARs. These features can help select a template and fill some of its non-spatial information slots. Finally, data from the invisible world representation can fill the spatial parameters of the classifier predicate PAR.

Since arm movements are represented as PARs, this design can use a planning process (like that of the AnimNL software) to transform these PARs into a 3D animation script (requirement 4). While the AnimNL's planning process turned invisible world PARs into animations of invisible objects, this planning process will turn classifier predicate PARs into an animation script controlling the movement of the signing character's arm as it produces a classifier predicate. (See Figure 2.)

Generating Multiple Classifier Predicates

Up until now, this paper has focused on generating a single classifier predicate from a single

English sentence, but in fact, the actual English-to-ASL translation problem is more complex. New challenges arise when generating several interrelated classifier predicates to describe a single scene. While specifying a system to generate a single predicate has been a natural starting point (and a first priority), it is important to consider how this architecture would need to be enhanced to handle the production of multiple classifier predicates. If these issues are not considered early in the development process, then software design decisions may be made that would make the MT system difficult to extend.

While the earlier sections of this paper may have suggested that there is always a correspondence between a single English input sentence and a single ASL classifier predicate output, in fact, several classifier predicates may be needed to convey the semantics of one English sentence (or vice versa). Even when the mapping is one-to-one, the classifier predicates may need to be rearranged during translation to reflect the scene organization or ASL conventions on how these predicates are sequenced or combined. For instance, when describing the arrangement of furniture in a room, signers generally sequence their description starting with items to one side of the doorway and then circling across the room back to the doorway again. An English description of a room may be significantly less spatially systematic in its ordering.

Multiple classifier predicates used to describe a single scene may also interact with and constrain one another. The selection of scale, perspective, and orientation of a scene chosen for the first classifier predicate will affect those that follow it. If decisions about the representation of the virtual reality scene are made without considering the requirements of the later classifier predicates, then output may be produced which arranges the elements of the scene in a non-fluent manner. Often the first English sentence describing a 3D scene may not contain enough detail to make all of the choices about the scene layout or perspective. A generation approach that considers the spatial information in adjacent (later) English input sentences prior to making such decisions could produce higher quality ASL output.

Another motivation for making generation decisions for groups of related classifier predicates is that the semantics of multiple classifier predicates may interact to produce emergent meaning. For example, one way to convey that an object is between two others in a scene is to use three classifier predicates: two to locate the elements on each side and then one for the entity in the middle. In isolation, these classifier predicates do not convey any idea of a spatial relationship, but in coordinated combination, this semantic effect is achieved.

Classifier Predicate Motifs

An MT system could handle the translation complexities discussed above by using sets of multi-classifier templates called *motifs*. Instead of immediately triggering one ASL classifier as each sentence of an English text is encountered, now the system will represent collections of multiple interrelated classifier predicate templates that can be used together to describe a scene. These collective structures would allow generation decisions to be made at the scene-level, thus decoupling individual English sentences from individual classifier predicates. The motif structure could decide how many classifiers must be used to communicate some block of spatial information and how to coordinate and arrange them.

A motif would serve as a set of deep generation rules or patterns for constructing a series of ASL classifier predicates in a specific semantic genre – e.g. movement of vehicles, giving directions, furniture arrangement, movements of walking people, etc. While this paper focuses on movement and location predicates, motifs can be imagined for size and shape specifiers (e.g. stripes or spots on clothing), instrument classifiers (e.g. using handtools), and others. Each motif would contain conditional rules for determining when it should be employed, that is, whether a particular English input text is within its genre. Just like the classifier predicate PAR templates in design 3, motifs could be triggered by features of the analyzed English text.[4]

Motifs would use planning rules to select and sequence their component predicates and to choose the best viewpoint, orientation, and scale for the entire scene. Having a separate motif for each genre would allow these planning rules to be specialized for how interrelated classifier predicates communicate spatial semantic information in a particular domain – possibly using genre-specific conventions as in the "furniture arrangement" example. Each motif could translate an English sentence according to its own guidelines; so, the system could translate the same input sentence differently based on the motif genre in which it occurred.

Implementation Issues

We can extend design 3 to generate multiple classifier predicates by adding a database of motif representations to be used in the PAR-planning process. In fact, these multi-predicate motifs could be represented as additional higher-level PAR templates. In the same way that a classifier predicate PAR can be hierarchically decomposed into sub-movements of the signer's arm (each represented by a lower-level PAR), analogously, a PAR representing a multi-predicate motif can be decomposed into PARs for individual classifier predicates. In design 3, English text features immediately triggered a single classifier predicate PAR; now, English features will trigger a PAR representing a motif. During planning, the motif PAR can use English text features and 3D invisible world data to decide how to expand its sub-actions – how to select and arrange the classifier predicates to express it.

Motifs are quite domain-specific in their implementation; so, questions can be raised as to what degree of linguistic coverage this design could achieve. This MT approach is certainly not meant to cover all English input sentences – only those that should be translated as classifier predicates. While domain-specificity can sometimes make an MT approach impractical to use, this design is meant to be embedded within a complete (possibly existing) MT system for English-to-ASL that uses traditional MT technologies to handle the majority of English inputs. Because these

[4] A stochastic motif genre-identifier could also be induced from statistical analyses of English texts known to produce a certain type of classifier predicate translation.

other MT processing pathways would be available, this design can focus on linguistic depth, rather than breadth.

With the linguistic coverage of the initial system as a baseline, the addition of this design would improve the coverage incrementally by bringing additional genres (domains) of classifier predicate expressions into the system's ASL repertoire as new motifs are implemented. The non-classifier translation pathways of the MT system would handle those spatial sentences still outside of the motif coverage. The other pathways would likely produce an overly English-like form of signing for these spatial sentences: a less desirable but somewhat useful result.

Relating Motifs to ASL Linguistic Models

The previously discussed linguistic models did not include a level of representation analogous to a motif because these models were focusing on a different part of the classifier predicate generation problem. Only after a signer has decided what spatial information to communicate (content selection) and how to sequence its presentation (propositional ordering) do these models describe how to build an individual classifier predicate (surface generation). They account for how humans produce single classifier predicate expressions – not how they plan the elements of an entire scene.

Linguistic models that do explain how human signers conceptualize 3D scenes also do not use a motif-analogous representation. Here, the reason may be that the generation task for a human is significantly different than the translation task for a computer. For example, Liddell (2003a) discusses how signers could plan a 3D scene and use multiple interrelated classifier predicates to describe it, but his model relies on the human ASL signers' rich mental visualization of objects in a 3D space and their ability to map (or "blend") these locations to the physical signing space. In a translation setting, the mental 3D visualization of the English speaker is not available; the English text is the only source of information about the scene. Because English generally includes less spatial detail than ASL when describing 3D space, both MT systems and human ASL interpreters are faced with the problem of understanding the English description and reconstructing the scene when producing classifier predicates.[5] Although not as robust as a human ASL interpreter, the AnimNL software can help this MT system create a 3D representation from the English text. But we are still left with the task of interpreting the English text for semantic and discourse cues to help guide our selection of classifier predicates to express this 3D scene. Therefore, motifs are triggered and informed by features from the analysis of the English text.

As a final linguistic concern, it is useful to consider whether the addition of motifs (that use 3D data) to design 3 has placed this system in further conflict with the polymorphemic model (Supalla, 1978, 1982, 1986). While this may initially appear to be the case, the addition of motifs is actually neutral with respect to this model. The model claims that an individual classifier predicate is composed from discrete morphemes, but it does not preclude the human signer from using mental 3D visualization of the scene during the deeper generation processes (those which overlap with the work of motifs). So, the point where the model diverges with this approach is the same as where it diverged from the original design 3 – when 3D data is used to fill the parameters of the classifier predicate PAR. This surface generation stage produces the non-categorical movements and locations of the classifier predicate output.

Discussion

Advantages of Virtual Reality

The 3D representation in this design allows it to consider spatial information when making generation decisions. Not only does this help make the generation of individual classifier predicates possible, but it also allows the system to potentially consider factors like spatial layout or visual salience when making deeper generation choices inside motifs – something a system without a 3D representation could never do.

This virtual reality representation for the space used by ASL classifier predicates may also be a basis for transcribing or recording these ASL phenomena electronically. A listing of the 3D objects currently in the invisible world with their properties/coordinates and a fully specified/planned arm movement PAR could be used to record a classifier predicate performance of a human signer. This approach would record more movement detail than classifier predicate glosses used in the linguistic literature, which merely describe the motion in English words and the handshape used. It would also be more informative than a simple movement annotation since it could store its non-spatial semantics (the semantic features that triggered the movement template), its spatial semantics (the locations of the 3D objects in the scene which it is describing), and the identities of those objects (what discourse entities are they representing). This additional information would likely be of interest to researchers studying these phenomena or building MT systems to handle them.

The 3D representation also allows this system to address ASL phenomena aside from classifier predicates in novel and richer ways. One example is the non-topological use of the ASL signing space to store locations for pronominal reference or agreement (Neidle et al., 2000). These locations could be modeled as special objects in the invisible world. The layout, management, and manipulation of these pronominal reference locations (or "tokens") is a non-trivial problem (Liddell, 2003a), which would benefit from the rich space provided by the virtual reality representation. If an ASL discourse model were managing a list of entities under discussion, then it could rely on the virtual reality representation to handle the graphical and spatial details of where these "tokens" are located and how to produce the "pointing" arm movements to refer to them.

The virtual reality representation could also facilitate the production of pronominal reference to entities that are "present" around the signing character. For instance, the character may be embedded in an application where it needed to refer to "visible" objects around it in the 3D virtual reality space or to computer screen elements on a surrounding user-interface. To make pronominal reference to an object in the visible 3D virtual

[5] And neither is perfect at this task.

reality space, a copy of this object could be made inside of the signing character's invisible world model. Then this invisible world copy could be treated like a "token" by the generation system, and pronominal references to this location could be made in the same way as for the "non-present" objects above. If the 3D object changed location during the signing performance, then its invisible world "token" counterpart can be repositioned correspondingly.

The AnimNL software makes use of sophisticated human characters that can be part of the scenes being controlled by the English text. These virtual humans possess many skills that would make them excellent ASL signers for this project: they can gaze in specific directions, make facial expressions useful for ASL grammatical features, point at objects in their surroundings, and move their hand to locations in space in a fluid and anatomically natural manner (Badler et al., 2000; Bindiganavale et al., 2000). When passed a minimal number of parameters, they can plan the animation and movement details needed to perform these linguistically useful actions. If one of these virtual humans served as the signing character, as one did for (Zhao et al., 2000), then the same graphics software would control both the invisible world representation and the ASL-signing character, thus simplifying the implementation of the MT system.

Current Work

Currently, this project is finishing the specification of both the classifier predicate generation design and a multi-pathway machine translation architecture in which it could be situated (Huenerfauth, 2004). Other research topics include: defining evaluation metrics for an MT system that produces ASL animation containing classifier predicates, developing PAR-compatible ASL syntactic representations that can record non-manual signals, and specifying ASL morphological or phonological representations that can be integrated with the PAR-based animation framework.

Acknowledgements

I would like to thank my advisors Mitch Marcus and Martha Palmer for their guidance, discussion, and revisions during the preparation of this work.

References

Bindiganavale, R., Schuler, W., Allbeck, J., Badler, N., Joshi, A., and Palmer, M. (2000). Dynamically Altering Agent Behaviors Using Natural Language Instructions. 4th International Conference on Autonomous Agents.

Badler, N., Bindiganavale, R., Allbeck, J., Schuler, W., Zhao, L., Lee, S., Shin, H., and Palmer, M. (2000). Parameterized Action Representation and Natural Language Instructions for Dynamic Behavior Modification of Embodied Agents. AAAI Spring Symposium.

DeMatteo, A. (1977). Visual Analogy and the Visual Analogues in American Sign Language. In Lynn Friedman (ed.). *On the Other Hand: New Perspectives on American Sign Language.* (pp 109-136). New York: Academic Press.

Holt, J. (1991). Demographic, Stanford Achievement Test - 8th Edition for Deaf and Hard of Hearing Students: Reading Comprehension Subgroup Results.

Huenerfauth, M. (2003). A Survey and Critique of American Sign Language Natural Language Generation and Machine Translation Systems. Technical Report MS-CIS-03-32, Computer and Information Science, University of Pennsylvania.

Huenerfauth, M. (2004). A Multi-Path Architecture for Machine Translation of English Text into American Sign Language Animation. In the Proceedings of the Student Workshop of the Human Language Technologies conference / North American chapter of the Association for Computational Linguistics annual meeting (HLT/NAACL 2004), Boston, MA.

Kipper, K., Snyder, B., Palmer, M. (2004). "Extending a Verb-lexicon Using a Semantically Annotated Corpus," In the Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC-04).

Liddell, S. (2003a). *Grammar, Gesture, and Meaning in American Sign Language.* UK: Cambridge University Press.

Liddell, S. (2003b). Sources of Meaning in ASL Classifier Predicates. In Karen Emmorey (ed.). *Perspectives on Classifier Constructions in Sign Languages.* Workshop on Classifier Constructions, La Jolla, San Diego, California.

Morford, J., and MacFarlane, J. (2003). "Frequency Characteristics of American Sign Language." Sign Language Studies, 3:2.

Neidle, C., Kegl, D., MacLaughlin, D., Bahan, B., and Lee, R.G. (2000). *The Syntax of American Sign Language: Functional Categories and Hierarchical Structure.* Cambridge, MA: The MIT Press.

Sáfár, É., and Marshall, I. (2001). The Architecture of an English-Text-to-Sign-Languages Translation System. In G. Angelova (ed.) *Recent Advances in Natural Language Processing (RANLP),* (pp. 223-228). Tzigov Chark, Bulgaria.

Schuler, W. (2003). Using model-theoretic semantic interpretation to guide statistical parsing and word recognition in a spoken language interface. Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics (ACL'03), Sapporo, Japan.

Speers, d'A. (2001). Representation of American Sign Language for Machine Translation. PhD Dissertation, Department of Linguistics, Georgetown University.

Supalla, T. (1978). Morphology of Verbs of Motion and Location. In F. Caccamise and D. Hicks (eds). *Proceedings of the Second National Symposium on Sign Language Research and Teaching.* (pp. 27-45). Silver Spring, MD: National Association for the Deaf.

Supalla, T. (1982). Structure and Acquisition of Verbs of Motion and Location in American Sign Language. Ph.D. Dissertation, University of California, San Diego.

Supalla, T. (1986). The Classifier System in American Sign Language. In C. Craig (ed.) *Noun Phrases and Categorization, Typological Studies in Language, 7.* (pp. 181-214). Philadelphia: John Benjamins.

Zhao, L., Kipper, K., Schuler, W., Vogler, C., Badler, N., and Palmer, M. (2000). A Machine Translation System from English to American Sign Language. Association for Machine Translation in the Americas.

A SIGN MATCHING TECHNIQUE TO SUPPORT SEARCHES IN SIGN LANGUAGE TEXTS

Antônio Carlos da Rocha Costa, Graçaliz Pereira Dimuro, Juliano Baldez de Freitas

ESIN/UCPel – Escola de Informática
Universidade Católica de Pelotas
{rocha,liz,jubafreitas}@atlas.ucpel.tche.br

Abstract

This paper presents a technique for matching two signs written in the SignWriting system. We have defined such technique to support procedures for searching in sign language texts that were written in that writing system. Given the graphical nature of SignWriting, a graphical pattern matching method is needed, which can deal in controlled ways with the small graphical variations writers can introduce in the graphical forms of the signs, when they write them. The technique we present builds on a so-called degree of graphical similarity between signs, allowing for a sort of "fuzzy" graphical pattern matching procedure for written signs.

1. Introduction

For the most part, software for processing sign language texts and databases have started to be developed only recently, simultaneously with the spreading of interest in SWML (Costa, 2003) among software developers concerned with the SignWriting (Sutton, a; Sutton, c) system. Obviously, an important and critical operation needed for such sign language processors is that of searching signs in sign language texts.

This paper presents a technique for matching two signs written in the SignWriting system. We have defined such technique to support procedures for searching sign language texts that were written in that writing system. Given the graphical nature of SignWriting, a graphical pattern matching method is needed, which can deal in controlled ways with the small graphical variations writers can introduce in the graphical forms of signs when they write them. The technique we present builds on a so-called degree of graphical similarity between signs, allowing for a sort of "fuzzy" graphical pattern matching procedure for written signs.

The paper is organized as follows. In section 2., we review aspects of sign languages related to the problem of having them written in some notation, and summarize the main features of the SignWriting system. Section 3. summarizes the work done on SWML and its importance for the development of software for processing SignWriting texts and databases. Section 4. presents the main contribution of the paper, namely, the sign matching technique designed to support procedures for searching in sign language texts. Section 5. brings the Conclusion. The sample signs presented in the paper are from the Brazilian sign language LIBRAS (Linguagem Brasileira de Sinais).

2. Sign languages and the SignWriting system

Along history, no writing system has been widely established for sign languages, so that such languages have always been used only for face-to-face communication.

Since Stokoe, in the 1960's, first recognized that sign languages are full natural languages, in the same sense that oral languages are, some notation systems for sign languages have been proposed. Stokoe himself introduced one such notation system (W. C. Stokoe and Croneberg, 1976). HamNosys (Hanke,) was another proposal. Both were conceived as technical tools for registering linguistic features of sign languages (handshapes, movements, articulation points, etc.).

SignWriting is also a proposed system for writing sign languages (Sutton, a). Contrary to the other systems, however, which were proposed mainly as tools for technical linguistic work, SignWriting was proposed as tool for daily use, by common (Deaf) people (Sutton, b).

3. SignWriting and SWML

Both the Stokoe system and HamNoSys are based on a linear representation of signs, using special characters for such purpose. SignWriting is based on graphical, bi-dimensional representations, using graphical symbols.

This way, the former systems can easily be encoded in computers in a linear way, by simply assigning numeric codes to each special character, and the technique for searching signs in texts written with such systems should be straight forward to develop.

SignWriting, on the other hand, requires that, besides the numeric encoding of each symbol, the computer representation of a sign keeps the information concerning the relative position of each symbol in the bi-dimensional area occupied by the representation of the sign (this complicates the searching procedure, as is shown below).

The SignWriter program (Sutton et al., 1995), the first computer editor for sign languages, defined such an encoding for SignWriting. That encoding was a binary encoding, created specifically for the needs of that program.

SWML (Costa, 2003) is a proposal for a general encoding format for SignWriting documents, using XML (**?**). It builds on the encoding used by the SignWriter program, presenting it in a fashion the makes such encoding available for use in all kinds of computer applications of SignWriting (document storage and retrieval, on-line dictionaries, computer interpretation and generation of sign languages, etc.). The `SW-Edit` program (Torchelsen et al., 2002) fully relies on `SWML` to store `SignWriting`-based sign language texts. `SignWriting` and `SWML` were proposed (Costa and Dimuro, 2002; Costa and Dimuro, 2003) as foundations for Sign Language Processing, the transposition of the methods and techniques of Natural Language Processing and Computational Linguistics, that have long been developed for oral language texts, to sign language texts.

The rest of this paper tackles one of the simplest operation one can do on a sign language document, namely, searching for a specific sign.

4. Matching Written Signs

There is a particular problem that has to be solved to allow sound searching procedures for sign languages files written in SignWriting, namely, to define a way of dealing with the small graphical variations that writers can introuce in the forms of the signs, when they write them.

The `SignWriting` system distinguishes explicitly some graphical properties of the symbols of a sign, like rotation and flop, for example, but does not distinguish tiny variations due to vertical and/or horizontal displacements of symbols within the sign, because such values are allowed to vary along the whole range of available positions within a sign box (as opposed to, e.g., rotation, which can only assume a small set of possible discrete values). The consequence of having such a "continuous" set of possible positions of symbols within a sign_box is that one lacks a clear geometric definition for the similarity between two signs, if they differ only with respect to the positions of their corresponding symbols.

The solution we have found to that problem is to allow the user to control the criteria to be used for judging on the degree of similarity of two signs by giving him a means to define a "fuzzy" correspondence between the component symbols of the two signs. The resulting matching procedure guarantees that two corresponding symbols have the same symbol type, rotation and flop, but allows them to have the (user specified) degree of variation on their relative positions within the respective signs instances. This kind of similarity between two signs is formalized in this section as a parameterized, reflexive and symmetric relation, that we call *sign similarity relation.*

4.1. Basic Geometric Features of Symbols and Signs

Initially, we formalize the basic geometric information concerning `SignWriting` symbols and signs.

Definition 1 *A* symbol s *is defined as a tuple* $s = (c, n, f, v)$*, where the values of* c, n, f *and* v *vary according to the symbol set being used, and:*

(i) *c is the category number (not available in symbol sets previous to the SSS-2002 symbol set (Sutton, c); use $c = 0$ in such cases),*

(ii) *n is the shape number (within the symbol's category),*

(iii) *v is the symbol variation (a complimentary information distinguishing symbols by features like, e.g., if the index finger is curved or not, in the symbol for the index handshape),*

(iv) *f is the filling information (encoding, e.g., palm orientation, in a symbol for a hand).*

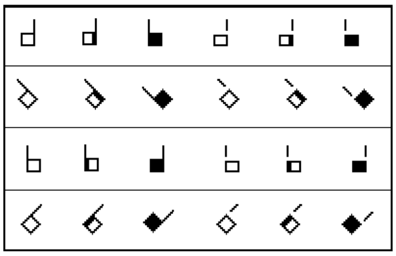


Figure 1: The group $G_{0/0}$ of symbols called *index*, and some of its rotated and flopped elements.

A set of symbols having the same symbol category and shape (c, n) and differing only in their filling or variation information, is called a *symbol group*, denoted by $G_{c/n}$. For each symbol group $G_{c/n}$ there is a so-called *basic symbol*, denoted by $s_{c/n}$, for which $f = 0$ and $v = 0$, so that $s_{c/n} = (c, n, 0, 0)$.

Definition 2 *An* oriented symbol ***S*** *is defined as a tuple* $\boldsymbol{S} = (s, r, fp)$*, where:*

(i) *s is a symbol of any symbol group $G_{c/n}$,*

(ii) *r indicates the (counter clockwise) rotation operation applied to s, relative to the basic symbol $s_{c/n}$ of the symbol group $G_{c/n}$ (the rotation is given in intervals of 45 degrees, for all symbols sets available up to now), and*

(iii) *fp, called flop, is a Boolean value indicating if the symbol s is vertically mirrored or not, relative to the basic symbol $s_{c/n}$ of the symbol group $G_{c/n}$.*

Example 1 *The symbol group called* index, *denoted by* $G_{0/0}$*, whose symbols, with category* $c = 0$ *and shape* $n = 0$*, represent hands with index finger straight up and closed fist, is shown in Figure 1. Each symbol s in the group $G_{0/0}$ is a tuple $s = (0, 0, 0, f)$, with variation $v = 0$ and fill $f = 0, 1, ..., 5$ (from left to right in the figure). The oriented symbols in the first row have the basic orientation (no rotations, no flop) and are given by tuples of the form $\boldsymbol{S} = (s, 0, 0)$. Each different fill information is represented by a different color fill in the symbol, indicating a different palm orientation, starting with the palm oriented towards the signer's face. In the second row, a rotation of 45 degrees was applied to each symbol, and the oriented symbols in that line are thus given by $\boldsymbol{S} = (s, 1, 0)$. In the third and fourth rows, representing the left hand, there are flopped symbols, given by $\boldsymbol{S} = (s, 0, 1)$ (with no rotations) and $\boldsymbol{S} = (s, 7, 1)$ (with rotations of 315 degrees).*

Definition 3 **(i)** *A* symbol box *is the least box that contains a symbol, defined as the 4-uple* $sb =$

(x, y, w_{sb}, h_{sb})*, where* x *and* y *are, respectively, the horizontal and vertical coordinates of the upper left corner of the symbol box (relative to the upper left corner of the sign box containing the symbol box — see item* (iv)*),* w_{sb} *is its width and* h_{sb} *is its height;*

(ii) *A* symbol instance*, that is, an occurrence of an oriented symbol within a sign, is defined as a pair* $\boldsymbol{Si} = (\boldsymbol{S}; sb)$*, where* $\boldsymbol{S} = (s, r, fl)$ *is an oriented symbol and* sb *is its symbol box;*

(iii) *A* sign*, denoted by* ***Sg****, is a finite set of symbol instances;*

(iv) *A* sign box *is a box that contains a sign, defined as a pair* $Sgb = (w_{Sgb}, h_{Sgb})$*, where* w_{Sgb} *is the box width and* h_{Sgb} *is the box height;*

(v) *A* sign instance *is defined as a tuple* $\boldsymbol{Sgi} = (\boldsymbol{Sg}; Sgb; p)$*, representing a sign* ***Sg*** *together with a sign box* Sgb *that contains it, and an index* p *indicating the position of the sign instance within the sign sequence (sign phrase) to which it belongs .*

All the definitions presented above are reflected in the SWML format. Note, in particular, that as defined above, sign boxes (and consequently, sign instances) have no coordinate information. This is so because sign language texts should be conceived essentially as strings of signs, with no particular formatting information included in them.

SWML, however, defines the notions of *document*, *page*, *line* and *cell*, so that sign instances can be put into cells, sequences of cells organized into lines, sequences of lines into pages, and sequences of pages into documents, in order to support document rendering procedures (e.g., horizontal or vertical renderings). Note also that symbols don't have predefined sizes (width and height). Sizes are defined only for symbol instances, through their symbol boxes. This allows for scalable symbol sets (e.g., in the SVG format (**?**)).

Example 2 *The* SWML *representation of the LIBRAS sign for* IDEA *(written as in Figure 2) is:*

```
<signbox>
 <symb x="46" y="37" x-flop="0" y-flop="0"
    color="0,0,0">
   <category>04</category>
   <group>02</group>
   <symbnum>001</symbnum>
   <variation>01</variation>
   <fill>01</fill>
   <rotation>04</rotation>
 </symb>
 <symb x="81" y="48" x-flop="0" y-flop="0"
    color="0,0,0">
   <category>01</category>
   <group>01</group>
   <symbnum>001</symbnum>
   <variation>01</variation>
   <fill>02</fill>
   <rotation>02</rotation>
 </symb>
 <symb x="62" y="18" x-flop="0" y-flop="0"
    color="0,0,0">
```

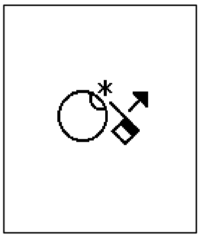


Figure 2: A way to write the LIBRAS sign for IDEA.

```
   <category>02</category>
   <group>01</group>
   <symbnum>001</symbnum>
   <variation>01</variation>
   <fill>01</fill>
   <rotation>01</rotation>
 </symb>
 <symb x="99" y="31" x-flop="0" y-flop="1"
    color="0,0,0">
   <category>02</category>
   <group>05</group>
   <symbnum>001</symbnum>
   <variation>01</variation>
   <fill>01</fill>
   <rotation>02</rotation>
 </symb>
</signbox>
```

4.2. The Sign Similarity Relation

The *sign similarity relation* is a parameterized, reflexive, symmetric and non transitive relation, introduced here to formalize the approximate similarity between two sign instances, and to provide for the construction of matching procedures for signs and sign language expressions.

The sign similarity relation has to embody an admissible difference in the positions of corresponding symbol instances within the two sign instances that it relates, taking into account a measure of significance for this difference, as determined by the user. The admissible differences in the positions of corresponding symbol instances are expressed in terms of percentages of some reference sizes, by a so-called *minimum degree of correspondence*, denoted by ε.

The reference sizes may be given either explicitly (e.g., 10 pixels) or implicitly (e.g., as the height and width of some symbol instance, chosen for that purpose among the symbols of the symbol set).

More over, the admissible difference in the corresponding positions of the corresponding symbols may be calculated in two ways:

- with respect to their *absolute* positions within the sign boxes to which they belong
- with respect to their positions *relative* to some reference symbol, known to be instantiated in each of the signs being compared

The *absolute* way of calculating the admissible differences is simpler, but the *relative* way allows the establishment of the similarity between a sign and another deriving

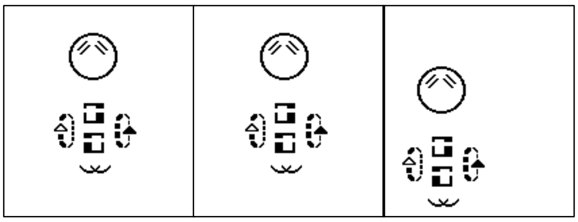


Figure 3: Similarity based on absolute and relative positions of the symbols (LIBRAS sign for YEAR).

from it just by a joint displacement of the symbols within the sign_box: e.g., in figure 3, the first sign instance would usually be judged similar only to the second instance, according to an absolute position based similarity relation, while it could also be judged similar to the third instance, according to the relative position based similarity relation.

We now define the sign similarity relation based on the absolute positions of the symbols.

Definition 4 *Let* $\boldsymbol{Si}_1 = (\boldsymbol{S}_1; sb_1)$ *and* $\boldsymbol{Si}_2 = (\boldsymbol{S}_2; sb_2)$ *be two symbol instances belonging to two different signs. Let their symbol boxes be given by* $sb_1 = (x_1, y_1, w_{sb1}, h_{sb1})$ *and* $sb_2 = (x_2, y_2, w_{sb2}, h_{sb2})$*, respectively. Then,* $\boldsymbol{Si}_1$ *and* $\boldsymbol{Si}_2$ *are said to* correspond to each other with at least degree ε*, and reference sizes* h_0 *and* w_0 *(for height and width), denoted by* $\boldsymbol{Si}_1 \approx^{\varepsilon}_{h_0,w_0} \boldsymbol{Si}_2$*, if and only if the following conditions hold:*

(i) Equality between the basic symbols*:*
$\boldsymbol{S}_1 = \boldsymbol{S}_2$ *(which implies* $w_{sb1} = w_{sb2}$ *and* $h_{sb1} = h_{sb2}$*),*

(ii) Admissible horizontal difference*:*
$|\frac{x_1 - x_2}{w_0}| \leq k$

(iii) Admissible vertical difference*:*
$|\frac{y_1 - y_2}{h_0}| \leq k$

where $k = \frac{100-\varepsilon}{100} \geq 0$.

Definition 5 *Let* $\boldsymbol{Sgi}_1 = (\boldsymbol{Sg}_1; Sgb_1; j_1)$ *and* $\boldsymbol{Sgi}_2 = (\boldsymbol{Sg}_2; Sgb_2; j_2)$ *be two sign instances.* $\boldsymbol{Sgi}_1$ *and* $\boldsymbol{Sgi}_2$ *are said to be* similar with at least degree ε*, relative to the absolute positions of their symbols, and reference sizes* h_0 *and* w_0*, if and only each symbol in a sign has one and only one corresponding symbol in the other sign, that is, there exists a bijection* $f : \boldsymbol{Sg}_1 \rightarrow \boldsymbol{Sg}_2$*, such that for each* $\boldsymbol{Si} \in \boldsymbol{Sg}_1$*,* $\boldsymbol{Si} \approx^{\varepsilon}_{h_0,w_0} f(\boldsymbol{Si})$.

Example 3 *Consider the three instances of the LIBRAS sign* IDEA *which are in Figure 4. Observe that each such sign instance contains an instance of the symbol* index *which differs in its coordinates from the corresponding* index *symbol instance of the other sign instances (all other symbol instances match exactly their correspondents). Consider a situation where a user is searching for that sign* IDEA *in a text. Suppose he writes the first sign instance as the sign to be searched and that only the two other instances are present in the text. The later two instances have some degree of similarity with the first sign instance. In spite of this fact, they are graphically different from the first instance, in a strict sense. They may all be considered to represent the same sign, or not, depending on the minimum degree of similarity required by the user for the results of the matching procedure. If the user specifies an intermediate degree of similarity, the second instance would match the first, while the third instance would not (the hand is too low in comparison with its position in the first sign instance). If the user specifies a low degree of similarity, all instances would match. If the user required 100% of similarity, no instance would match. The total degree of similarity (*$\varepsilon = 100\%$*) requires that no difference be admitted between the two sign instances being compared.*

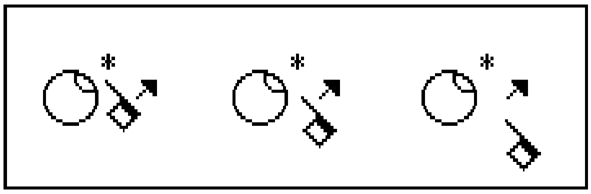


Figure 4: Three (possible) instances of the LIBRAS sign *IDEA*.

The basic similarity relation defined above does not take into account some important (and frequent) exceptions. Such exceptions are mainly related to symbols like the *arrow* symbol (encountered, e.g., in the LIBRAS sign IDEA), whose position within the sign is, in general, not critical (see Figure 5). Such symbols have most of their meaning completely encoded in their shapes and transformations, and the place where they are put in the sign boxes is essentially irrelevant. For instance, the *arrow* symbol in the sign for IDEA means that the right hand moves in the horizontal plane, in the indicated direction, and this information is the same, wherever the *arrow* is placed in the sign box. In such cases, the relative position of the symbol within the sign box is not important. In the examples of the Figure 5, even if a rigorous or a total degree of similarity is required, the matching process should find that those three sign instances are similar. On the other hand, for symbols like the *asterisk*, almost no variation of the its position should be allowed, since it indicates a position where two components of the sign (e.g., head, hands, etc.) touch each other when the sign is performed, and even small degrees of variations may imply linguistically relevant differences between the signs.

Other reasonable definitions for the sign similarity relation could be given such as, for instance, the one already mentioned, of taking the positions of the symbols relatively to a reference symbol, known to occur on both the sign instances that are being compared. Even coarser relations could be defined, and possibly considered useful, e.g., one defining the admissible differences on the basis of the absolute coordinates of the very symbols being compared.

4.3. Search Procedures for Sign Texts

`SWML`, as currently defined, already has all information

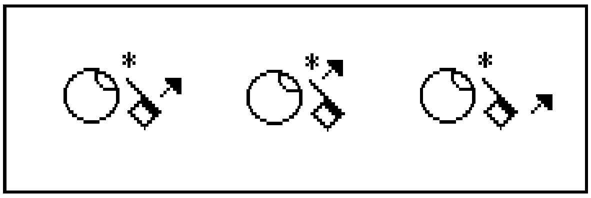


Figure 5: Three (guaranteed) instances of the LIBRAS sign *IDEA*.

needed to allow for a *sign matching procedure* based on the sign similarity relation defined here. The special treatment of symbols whose meanings are not sensitive to the symbols' placements in the signs is to be embedded in the matching process, requiring from `SWML` only that it identifies symbol instances completely, which it perfectly does. On the basis of such sign matching procedure, a procedure to search for signs in sign language texts can be easily defined, in a straightforward way.

5. Conclusion

In this paper, we have shown that searching for signs in sign language texts written in `SignWriting` is a straight forward matter. The only slightly tricky part of the searching procedure is in the operation of matching two signs, which should allow for small differences in the positions of their corresponding symbol instances. Ideally, the size of the differences that are to be admitted in such correspondence tests should be specifiable by the user when he calls the search procedure, so that he can have full control over the desired degree of similarity of the signs being compared.

Acknowledgements

This work was partially supported by CNPq and FAPERGS.

6. References

A. C. R. Costa and G. P. Dimuro. 2002. `SignWriting`-based sign language processing. In I. Wachsmuth and T. Sowa, editors, *Gesture and Sign Language in Human-Computer Interaction*, pages 202–05, Berlin. Springer-Verlag.

A. C. R. Costa and G. P. Dimuro. 2003. `SignWriting` and `SWML`: Paving the way to sign language processing. In O. Streiter, editor, *Traitement Automatique des Langues de Signes, Workshop on Minority Languages, Batz-sur-Mer, June 11-14, 2003*.

A. C. R. Costa. 2003. The `SWML` site. Located at:. `http://swml.ucpel.tche.br`.

T. Hanke. Hamnosys - the hamburg notation system. Located at:.

V. Sutton. Lessons in signwriting – textbook and workbook. Available on-line at:. (Sutton, c).

V. Sutton. The signwriting history. Available on-line at:. (Sutton, c).

V. Sutton. The `SignWriting` site. Located at:. `http://www.signwriting.org`.

V. Sutton, , and R. Gleaves. 1995. *SignWriter – The world's first sign language processor*. Center for Sutton Movement Writing, La Jolla.

R. P. Torchelsen, A. C. R. Costa, and G. P. Dimuro. 2002. Editor para textos escritos em signwriting. In *5th. Simposium on Human Factors in Computer Systems, IHC 2003*, pages 363–66, Fortaleza. SBC.

D. C. Casterline W. C. Stokoe and C. G. Croneberg. 1976. *A Dictionary of American Sign Language on Linguistic Principles*. Linstok Press, Silver Spring.

A Practical Writing System for Sign Languages

Angel Herrero
University of Alicante
Ap.Correos 99.-E-03080 Alicante (Spain)
angel.herrero@ua.es

Abstract

This paper discusses the problems involved in writing sign languages and explains the solutions offered by the Alphabetic Writing System (*Sistema de Escritura Alfabética,* S.E.A.) developed at the University of Alicante in Spain. We will ponder the syllabic nature of glottographic or phonetically-based writing systems, and will compare practical phonological knowledge of writing with notions of syllables and sequence. Taking advantage of the ideas of sequentiality contributed by the phonology of sign languages, we will propose a sequential writing model that can represent signers' practical phonological knowledge.

1. Sign Languages and Writing

Except for semasiographic systems, such as the *winter counts* of the Dakota people, and visual instructions for the use of certain machines, which "state ideas directly", all writing systems are glottographic (Sampson, 1997: 42). In other words, "they […] use visible marks to represent forms of a spoken language". Writing systems that had initially been considered pictographic, such as Egyptian hieroglyphics, Chinese writing, Mayan glyphs, or the Easter Island tablets, were later shown to be glottographic, or "true writing", as underlined by the greatest scholar of writing systems, Thomas Barthel.

Ever since it was discovered by Sumerian culture, alphabetic writing has been based on syllables, involving a phonological analysis of the chain that bases representation on the different components of each syllable: consonants and vowels. Other glottographic writing systems, known as logographical writing systems, are based on significant parts of words, or morphemes. This is the case of Chinese for example, although in this case the significant parts of the words, the morphemes, generally coincide with syllables, meaning that logographic writing may also be considered syllabic. There are also cases of 'motivated' logographic writing systems, such as the phonological-featural alphabet of Korean Hangul. However, in this phonological-featural alphabet also, based on infra-phonemic elements, "the essential graphic distinction is between vowels and consonants" (Sampson, 1997:179). In practice, different writing systems can be combined, as we do when we use morphological symbols such as numbers or percentage symbols, present on all keyboards, in alphabetical texts.

The distinction between Consonant and Vowel has proven to be an excellent criterion for phonological representation: it is immensely practical, as it represents the syllable at the same time. In other words, and this is the essential idea of our proposed writing system, **Consonants and Vowels are represented as stages of articulation.** Non-segmental phonology, specifically feature-geometrical phonology (Clements, 1985) or Prosodic Phonology, on which the most complete model of ASL phonology, devised by Brentari (1998) is based, have resolved the CV difference in other minor differences, so that V or C is a relative question, arising from the assignation of features; the notions of V or C can be replaced by the notion of auto segment, or even by a phonological rule, thereby giving a more explicative model for certain phenomena such as tone, vocalic harmony or the vocalic morphology of certain oral languages. However, it has to be said that tonal languages and others that have been put forward to justify a non-segmental conceit in prosodic phonology (Venda, Turkish, Hebrew), are currently written in alphabetical, segmental writing.

From a scientific point of view, the practical phonology that gave rise to writing is full of imperfections, creating an unreal image of languages (Olson, 1991:285). However, this image has historically been identified with knowledge and culture, and writing, with all its imperfections, has become an irreplaceable practical skill for consigning knowledge. The reason for this is, doubtless, the way it represents the speech process.

Therefore, if sign languages, from the point of view of linguistic typology, are comparable to oral languages in many morphological and syntactical aspects, it would appear logical to extend this comparison to the syllable as the basic phonotactic unit of writing, although the concept of syllable is also currently questioned in non-linear phonology (Wilbur, 1990). If letters (characters) represent the kind (and stage) of articulation of the sounds in a syllable, so that the speaker can not only make the sounds but also distinguish the order in which they are produced, in sign languages (LSs) letters may also represent the kind (and stage) of manual articulation, and the order of the letters can represent the order of production of signs by the signer.

In this paper we will present a proposed writing system based on this possibility. Annotation systems currently used to transcribe signs, such as the *HamNoSys* system devised by Siegmund Prillwitz and his group at Hamburg University, or *SignWriting* devised by Valery Sutton at San Diego University, may not be processed as writing. *SignWriting* showed the very possibility of writing and is a historic contribution to the culture of the signing community, but the alphabetical writing system we present is based on a principle of phonological economy, while *SignWriting,* because of its openly visual nature, is based on simultaneity and the supposed analogical or iconic nature of the signs. The problems with alphabetical writing are precisely the advantages of *SignWriting*: the supposed simultaneity of the signs and their analogical nature, particularly obvious in non-manual expression. We will now see that the notion of simultaneity goes hand-in-hand with the notion of syllable and that they have compatible sequential processes.

2. Syllable, Sequence and Simultaneity

Although the current phonology of sign languages still suffers from many problems, as can be seen from the different phonological models that have been devised one after the other in recent years (Liddell, 1984, 1989, 1990; Sandler, 1989; Perlmuter,1988; Brentari ,1998, 2002), there is still sufficient consensus, in our opinion, to justify a proposed writing system that could be used as a skill, rather than a phonological model.

As we have pointed out, the basic unit of glottographic writing is the syllable, as this is the minimum unit in which sounds can be distinguished and combined. Accordingly, in spite of certain pending questions (such as the phonological interpretation of repetition and lengthening), the phonology of sign languages already gives a good idea of the phonological components of the syllable and its limits. It is also generally agreed that two successive movements, even when they are local, correspond to two syllables, and rules have been made for elision, epenthesis and gemination (Liddell, 1989). However, the main problem with these methods is that they continue to consider that, except for the movements, which, by definition, are sequential, the syllable is simultaneous.

In 1933 the vocal apparatus was filmed in operation for the first time, and the great linguist Roman Jakobson was very impressed by the result. In the first of his *Six leçons sur le son et le sens,* given in New York in 1942, he remembers the film and states (1988: 396) that when he saw it he understood that "the act of speaking is a continuous, uninterrupted movement… there are no position vs. transition sounds; they are all transition(...) Strictly from the point of view of articulation, the sequence of sounds does not exist. Instead of following each other, sounds link up with each other; and one sound, which our acoustic impression tells us comes after another, can be articulated simultaneously with it or even partially before it(...) It is not possible to classify, or even, I would say, to describe the different articulations accurately, without continuously asking what is the acoustic function of such and such motor action" .

Syllables are acoustic units determined by the level of merging and influence of vowels and consonants (Malmberg, 1955), which are, therefore, relative segments. Syllables are recognised by the transitions of a vowel or nucleus due to the effect of the consonant(s) of the syllable.

Thus, as its etymology indicates, the syllable is a paradigm of simultaneity. In written representation, we would point out that literate speakers recognise segments of this transition; a segmental sound is an articulation with stable parameters, insofar as there are changes between the sounds that allow us to identify them. Accordingly, the real effect of the operation is simultaneity, while segmentality is an operation of the mind, which I have described above as practical phonological knowledge, distinguishing between CV and types of both.

So what segments should be represented in writing an SL, in our case Spanish Sign Language (LSE)? The linguistics of sign languages was born with the discovery of its phonemes (Stokoe, 1960), initially called phonological 'aspects' or 'cheremes' and later, 'parameters', a term which has spread to most current phonological models (e.g. Brentari, 2002). Until the 80's, these constituents, which we believe should simply be considered phonemes, were seen as simultaneous with monosyllabic signs, i.e., syllables. A fourth parameter, Orientation, was added to the three proposed by Stokoe (1960): Location, Hand Shape and Movement, sometimes called the major parameters, and the difference between path movements and local movements was specified (Liddell, 1989). Additionally, the passive hand should be specified as the location L of the sign when it acts as such, with its own Q and O, or as an element of symmetry with the active hand. Lastly, our writing system represents possible contact with the body, C, as a specification of location. These are the constituents that we represent.

We are not going to deal here with the phonological or featural nature of these components, but briefly to justify their sequential representation and the use of the Hand Shape as the nucleus of the syllable, as the basis for an economical writing system.

2.1. Sequentiality

Several sequential models have been proposed since the 80's: Liddel (1982, 1989), Sandler (1986, 1989, 1990), Perlmutter (1988), Brentari's prosodic model (1998), etc. In this last one, Hand Shape, Location, Orientation and Movement are treated as types of (geometric) features, rather than segments. It considers that, "It is sufficient to make reference to distinctive features, in syllable initial and syllable final positions, and there is no support for any further internal segmental divisions... no intermediate segments are recognized by the signers". Moreover, Brentari (2002: 45) considered that simultaneity is a characteristic of sign languages, "Cs and Vs are realized at the same time in sign languages, rather than as temporally discrete units"; (2002:47): "If sign language Cs are properties of the IF tree and sign language, Vs are properties of the PF tree, the major difference between sign and spoken languages in this regard is that in sign languages IFs and PFs occur at the same time"

Liddel's model conceived of Hold and Movement as segments, so that its syllabic model consisted of a hold-movement-hold sequence; the Hand Shape and Orientation features, along with contact and Location L, formed part of specific tiers, represented as simultaneous. Sandler's model is also partially sequential, based on Location and Movement segments; this model also recognises the segmental nature of Q (Sandler, 1990:20 "hand shape is a distinct and temporally autonomous phonological element in ASL"). In our proposal, sequentiality will be extended to all the other parameters, although we insist that our aim is not to present a phonological model, but rather a model of written representation. This model, which we call the Alphabetical Writing System for Spanish Sign Language (Sistema de Escritura Alfabética de la Lengua de Signos Española - SEA.), is available in book form (Herrero and Alfaro, 1999; Herrero, 2003) and on the internet (cervantesvirtual.com/portal/signos); all we can do here is describe its essential elements in relation to the problems that practical phonology based on writing may raise when approaching theoretic phonology. The system has been successfully taught to several signers in a few weeks.

For our writing system, we start off by taking the basic sequence proposed by Sandler (in its turn a specification of the one proposed by Liddell): the Location-Movement sequence. There are several pairs of signs that show the sequential incidence of Movement:

AMOR ....LASTIMA

love *pity*
DIFÍCIL ....ANUNCIAR
difficult *to announce*
JUNTOS..........MESA
together *table*
LISTO ....SABER
clever *to know*
TELEFONO ...LLAMAR POR TELEFONO
telephone *to phone*
ARBOL...........BOSQUE
tree *forest*
SILLA ....SILLAS
chair *chairs*
MIRAR ....VER
to look *to see*
ARRIBA........ .MANDAR
up *to command*
LLAVE .... ESPADA
key *sword*
CASA .... CASA GRANDE
house *big house*
PROBAR...... ARADO
to try *plough*

Using this elemental sequence, which refers only to two phonemes or parameters, Location and Movement, the remaining parameters are written in the following order:

S L(.)QODF

Where

- S represents the left hand (as in ESCRIBIR , *to write*) or active two-handed signs (as in VIAJE, *journey*).
- The point (.) that may follow Location indicates that there is no contact with the part of the body taken as reference for signing (the temple, in TEORIA, *theory*)
- Hand Shape Q and Hand Shape Orientation follow after Location and before Movement
- Movement M is differentiated, as is normal in all phonological models, into Path Movement (D) and Local Movement (F), which are not obligatory, may be simultaneous and, when simultaneous, give rise to two syllables. The simultaneity of D and F will be represented by adding the direction feature to the F symbol, i.e., making a kind of D out of DF.
- Non-manual elements that accompany the signs will only be represented if they have morphological value (e.g., adverbial intensification, although most signers know lexical forms of representing this intensification; or simultaneous affirmation and negation).

Before going deeper into the writing system and giving examples, we would first like to make a few comments on the decisions that we have taken and that we have just summarised.

a) The initial writing of the passive hand when it acts as Location (but not in two-handed signs or as the moving hand) is justified by articulatory and perceptive reasons: while making the sign, the dominant hand addresses the *previously* moving passive hand (ESCRIBIR, *to write*; POR QUÉ, *why*; OBJETIVO, *aim*). As far as I know, this sequentiality has so far gone unnoticed.

b) We also consider it proper to represent active two-handed signing (symmetric, asymmetric and displaced symmetric signs) at the beginning for reasons of processing, as two-handedness affects the articulation of all the other components from the beginning.

c) We have already said that there is a general consensus as regards the Location-Movement sequence. The Hand Shape and Orientation components are represented between the two. On the one hand, it would appear obvious that what Movement does is to modify Location, in the case of Path Movement D, or Hand Shape Q and/or Orientation O in Local Movements; these components should be specified before M as they are a part of the Hold (in Liddell's model).

d) The LQO order is an interpretation of the articulation of bringing the hand from a part of the body or from the signing space with an articulation Q. The hand then remains in that Location with a certain Orientation and, in dynamic signs (most of them), carries out a movement.

e) The precedence of L over Q is clear when L is the passive hand. Another indication is given by the fact that when the sign is made in the mouth (SILENCIO, *silence*; ODIO, *hate*; ROJO, *red*) the position of the lips goes before Q, and when the sign is made with a non-manual component (DELGADO, *thin*), this component goes before Q. In general, this place is guessed "before" Q, as a root which Q will specify. As a matter of fact, the initial process of articulation in many signs is similar to an oral CV syllable, insofar as the articulation takes the Hand Shape of Q as that of the Vowel, while the occlusion occurs. We use the term 'occlusion' here in the sense of visual perception studies, as occlusion (interposition) of one object by another, in this case, the body by the hand (Kanisza, 1986: 283). What is not seen is not so much a mental representation as a 'found detail in an non-modal complementation, with clear functional effects on the perception of fragmented objects.

f) One last clarification regarding sequentiality: Movement, whether path or local, does not generally have a specified ending place. The sign does not necessarily stop in one place (IDEA, *idea*; ENFADADO, *angry*) and, if it does, does not do so in a lexical Location L (but rather in a precisely moved place), or with a Hand Shape Q or an orientation O other than those foreseen by M, these Locations, Hand Shapes and Orientations being moreover subject to strict constrictions. M consists precisely of leading to that end. Another thing is two successive movements (ESPADA, *sword*), or two phonological places (PADRE , *father*), which we consider disyllabic, but in monosyllabic signs the economy of the writing system makes it possible to end the sign in its movement. The incidence of certain Ms, specifically in local Fs, which modify Q and/or O, seems comparable to glides in oral languages. D movements, on the other hand, do not change Q and can be compared to consonants. The incidence of M is phonetically very varied.

We now give some arguments for considering Q the syllabic nucleus, and thus justify its being written in the centre of the syllable.

2.2. The Nuclear Character of Q

We agree with Brentari (1998: 313) that "the formal role of distinctive features, syllables, and segments as building blocks of a grammar with constraints is the same for signed and spoken languages, but the substantive definitions in both types of languages –those that are more phonetic and less grammatical- depend on conditions of naturalness in each modality", although we believe that the identity of the formal role should be translated as the difference between nucleus, onset and coda (or between onset and rhyme), which is immensely important, as far as writing is concerned. This is the difference on which the writing system is based, and, although the model is not the most scientifically suited for the phonological description of sign languages, as neither is it for oral languages (according to non-linear phonology), it may be applied to sign languages with similar criteria as to spoken languages. This opinion is defended by Wilbur (1990).

The following are the main reasons why we will consider Q the nucleus:

a) The nucleus is a necessary constituent of every syllable. Some phonologists have stated that the necessary, nuclear, constituent is Movement. Brentari (2002:44), for example: "regarding minimal word constraints, no sign is well formed unless it has a movement of some type", but, in Spanish Sign Language at least, there are fairly evident counter-examples of signs without M: one-handed signs such as OJO *(eye)*, ALTO *(tall)*, ANCHO *(wide);* and two-handed signs such as PELOTA *(ball)*, GAFAS *(glasses)*, CRUCIFIJO *(crucifix),* which neither have movement nor undergo an epenthesis of movement, as Brentari states. On the other hand, the only signs without Q are the non-manual signs (Dively, 2002). These signs are generally gestures (emblems, etc.), and have no lexical entity. When they act with related morphological value, they are represented at the end of the sign.

b) While Location or Movement can be reduced in rapid signing (IDEA, *idea*, can be signed in a slightly higher place, although not at the temple; or the movement of EMPEZAR, *to begin*, can be reduced to a slight, local waving movement), Hand Shape cannot usually be reduced.

c) We agree with Coulter (1990: 125) that stress is "the notion that greater articulatory effort is involved", i.e. as muscular tension, so that, according to Wilbur (1990: 99) "stressed signs were temporally shorter than unstressed". In prosodic phonological models, the nuclear nature of Movement means that it carries prosodic marks such as duration, but I believe that this is not the same as stress. In this regard, it is very significant that the emphasis on some signs normally made with binary repetition eliminates this repetition while tensing the articulation. We believe that our point of view is compatible with the well-known Straka Rule, "under the effect of reinforcing articulatory energy, consonants close and vowels open; on the contrary, under the effect of articulatory weakening, consonants open and vowels close" (Straka, 1963: 35)

d) Lastly, it should be noted that when Sign Languages are interpreted for deaf-blind people, they are reduced to Q, insofar as fingerspelling is a part of Sign Languages.

Considering Q the nucleus also resolves the problem of Hand Shape double behaviour in prosodic models. As regards this double behaviour Corina (2002: 91-92) has said, "that is, that hand shapes may be a constituent of the syllable nucleus or not" or, in other words (Corina, 2002: 94) "in instances when the hand shape changes, hand shape is functioning more like a vowel. In those signs with no change in hand shape, hand shape serves a more consonantal function". Brentari (2002:30) has also referred to this double status, "Depending on whether the posture of the hands remains constant throughout a sign – in which case the dynamic portion of the signs comes from path movement-or whether the posture of the hands changes while the other parameters are held constant, hand configurations can be thought as non nuclear (C) or nuclear (V) in a sign syllable". We could also ask about simultaneous changes in hand shape and path movement (as in COMPRENDER, *to understand*), which would involve a new treatment of hand shape. However, its unified treatment as a nucleus avoids these dysfunctions.

In our model, the components or phonemes of Location, Hand Shape, Orientation and Movement can be considered structurally or syntactically as the [Onset] [Rhyme (nucleus, coda)] elements of the syllable. This model has the asymmetrical conditions that characterise linguistic constructs, as regards syllabic structure (Carstairs-McCarthy, 2001).

3. Economy of the Writing System: Projection Model, Featural Elements and Rules for Simplification

When the Greeks imported Semitic writing, they gave the characters the Greek names closest in sound to their Semitic names (aleph / alpha), and adapted them to represent their own sounds (many of which, particularly learned words, were borrowed from Semitic languages). In sign languages, the alphabet may not be imported based on reasons of perceptive analogy, but on general semiotic values associated to different types of sounds.

Moreover, although the exact number of phonemes of each type (places on the body or in the signing space, hand shapes, types of orientation, types of movement) is not closed, at least in Spanish Sign Language, we know enough to propose a representation open to new symbols. What we do know is that the number of phonemes, understood like this, is clearly greater in Sign Languages than in spoken languages: 32 parts of the body, 10 parts of the signing space, 31 hand shapes, four orientations for each hand shape; as regards M, the number depends on the consideration of features. This complexity will be resolved by what means of what we call the projection model. In any case, this property of sign languages leads to a phoneme: morpheme ratio of almost 1:1.

The symbols (represented by consonants) for the parts of the signing space, orientation and direction of movement will be further specified by means of vowels, using a hand projection model which associates "up", "upwards" or "towards the signer's face" with the vowel "a" (which also symbolises the thumb); "down", "downwards" or "towards the listener's face" with the vowel "u" (which also symbolises the little finger); "left" towards the left" with the vowel "i" (which also symbolises the middle finger); "in front" or "forwards" with the vowel "e" (which also represents the index finger); "in the centre" or "backwards" with the vowel "o" (which also represents the hand shape that uses the five fingers); and "right" or "towards the right" with the symbol "y". This geometric model has been partially inspired by Friedman (1977).

These specifications are features that allow more analytical representation and easier reading. In the cases of Location, the sub-specification appears before the symbol for the place in space (the central longitudinal plane, symbolised by **l,** and the right longitudinal lateral, represented by the consonant **b**), so that **al** is the high part of the central plane (as in CIELO, *sky*); **el,** the frontal part of the same plane (as in TU, *you*); **ub,** the "low" part of the lateral plane (as in BAJO, *low*); **ab**, the high part of the lateral plane (CONFERENCIA, *conference*), etc.

In the case of the Orientation, after the consonant **m,** the sub-specifications use a first vowel to indicate the direction of the fingers of the hand (on the open palm); a second one, the orientation of the palm: natural orientation, or following on from the arm, which does not need to be represented and for which the first vowel is sufficient (as in CONFERENCIA, *conference*, **ma**; or in TU, *you,* **me**); orientation towards the signor or upwards (an **a** is added as in PASADO, *past,* **maa**; or in QUE, *what*, **mea**); orientation towards the listener or downwards (a **u** is added, as in COMPRENDER, *to understand*, **mau**; or in COGER, *to catch,* **meu**); and orientation towards the right or inversely to natural continuity with the arm (a **y** is added, as in SEPARARSE, *to separate*, **mey**). The same occurs with the other orientations for the direction of the fingers (**mi, mia, miu, miy; mu, mua, muu, muy** etc..).

In the case of Direction (D), the vowel added to the straight movement symbol (**w**) states the direction: **wa** is upwards, as in FUEGO *(fire);* **we,** forwards, as in CONFERENCIA *(conference)*; **wo**, backwards, as in COMPRENDER *(understand).* Curved directional movements are represented by a **c** followed by two vowels, one for direction and the other for curvature: **cea** would be a direction curve forwards curving upwards, as in DAR *(to give)*; **cya,** curve towards the right curving upwards, as in ARCO *(arch, bow)*, etc. These direction vowels are added directly to the local movement symbols when they are carried out with directional movement. Thus, the extension/flexion symbol **l** is followed by **o** to indicate extension/flexion moving backwards, as in COMPRENDER *(to understand),* which is why this word ends in **lo**; or a trembling movement, symbolised by **t**, is followed by **e** to indicate that it occurs in a forwards direction, as in BOSQUE *(forest)*, which is why this word ends in **te**. Local movements such as waving, beckoning and twisting, indicate the direction of their local movement with the respective vowels.

Some local movements are involved in symmetry (tapping or hitting between the two hands, linking, etc.) and, in this case, may be represented using the two-handed **s** symbol. For example, a symmetric tapping movement between the two hands, such as CONTACT *(contact),* will be symbolised by **sp**, where **p** is the symbol of the tapping F: a symmetric hitting movement, as in HIERRO *(iron),* is symbolised by **sx**, where **x** is the symbol of the hitting F, etc. The signs thereby will have a sequence as follows:

1. S (if it is two-handed) + indicators of the type of symmetry/QO of the passive hand
2. spacing
3. body consonant / vowel + l/b (Location)
4. optional point (Contact)
5. Q (Configuration)
6. m (Orientation) + orientation vowels
7. D consonant + direction vowel/s
8. F consonant/s + direction vowel/s

We have left the representation of Q for the end. To a certain extent, it is the easiest, insofar that every finger, except the ring finger, has a symbol, and it is easy to use diacritical symbols to indicate the features of flexion (´), union (`), contact (^) and link (¨), and to distinguish from the order of the fingers if the shape is open-handed (as in POLVO, *dust*) or close-fisted (as in MINUTO, *minute*).

The method presented here is completed with certain rules for the simplification of location and orientation, based on considering certain locations or orientations 'natural' and not symbolising them. Thereby, writing Spanish Sign Language becomes very easy.

We use the following two rules for the simplification of locations (not written):

a) Simplification of the **ol** central location of most two-handed signs.
b) Simplification of the lateral location (**ab, eb, ib, ob, ob, yb**) when the hand is in its natural position following on from the arm (**òma**, instead of **abòma**; **òmi**, instead of **ilòmi**, etc)

We use the following two rules for simplifying orientation:

a) Simplification of the orientation when the location is a part of the body, and the palm is oriented towards that location (e.g. **ynò**, rather than **yòmi**)
b) Simplification of the **me** orientation when the sign is made in **eb**, as occurs in many signs such as PISTOLA *(pistol)*, BASTON *(walking stick)*, REGULAR *(regular)*, etc.

Lastly, we simplify L and O by using only diacritical and numerical signs.

The possibility of alphabetical writing has been tested, writing all the signs contained in Spanish Sign Language dictionaries, particularly Pinero's dictionary (1989), and also in the translation of several texts, including poetry, and in teaching the method to groups of signers. However, as we have stated already, writing is not a reproduction of spoken language: it is a representation, a record, with its advantages and limitations, of the spontaneous act of signing. The lack of a prosodic representation of the writing of many oral languages is a limitation, particularly from the point of view of non-literate persons, although this limitation, related with the lack of context and the non-presence of the interlocutors, makes the written message very suitable for reflection, and very open to interpretation.

Writing signed spontaneous conversation generally involves adopting certain other symbols, particularly Location. According to Liddel (1990), in addition to the phonological places where said lexical signs are located (10 in the signing space and 32 on the body), there also exist anaphoric grammatical spaces and descriptive, analogical or topographical spaces, which copy the real situation of objects in real space, and are used in blended spaces in descriptions. There are no problems in applying the projection model to represent grammatical locations; descriptive locations may be represented by means of directional repetitions, but if this is not possible, they will have to be paraphrased by writing "to the left," "crossed," etc. This is also the case with many non-manual expressions describing modality, i.e. doubt, certainty, etc.

We now give the writing for certain Spanish Sign Language signs of different phonological composition. Disyllabic signs are written with a hyphen; L and/or O simplified using the rules mentioned above are written in brackets:

	2 hand	L C	Q	O	D F	disyllabic
amor *(love)*		**yn**	**i**	(mi)		
cauce *(course)*	**sm**	(ol)	**ò**	(me)	**se**	
rubio *(blond)*		**c**	**i**	(miu)	**zo**	
teoría *(theory)*		**t .**	**T**	(ma)	**wruhob**	
espada *(sword)*		(eb)	**aë**	**meu**	**cre**	**- we**
libro *(book)*	**sc**	(ol)	**ò**	(me)	**creb**	
sordo *(deaf)*		**r**	**e**	(mau)		**- v**
Portugal	**sm**	**pn**	**a**	**miu**	**zuy**	
ayer *(yesterday)*		**hm.**	**oa**	(maa)	**daheb**	
casi *(almost)*		(eb)	**aë**	**mea**	**grel**	
dar *(to give)*		**y .**	**aë**	**mo cea**		
China		**yn**	**e**	(mo)	**zy**	**- zu**
bilinguismo *(bilingualism)*	**so'ami**		**ei**	**mau**	**wu gre**	

Acknowledgements

This paper has been possible thanks to the economical support of Ministerio de Ciencia y Tecnología (Research Proyect nº BFF2002-010016).

References

Brentari, D. (1998) A Prosodic Model of Sign Language Phonology. New York: The MIT Press

Brentari, D. (2002): Modality differences in sign language phonology and Morphophonemics, in Meier, Richard P.; Cormier, Kearsy; and Quinto-Pozos, David (eds.), pp. 35-64

Carstairs-McCarthy, A. (2001) ASL 'syllables' and language evolution: a response to Uriagereka, in Language 77, 343-349

Clements, G.N. (1985) The geometry of phonological features. Phonology Yearbook 2: 225-252

Corina, D. & Sandler W. (1993) On the nature of phonological structure in sign language. Phonology 10: 165-207

Corina, D.P. and Hildebrandt, U.C.(2002): Psycholinguistic investigations of phonological structure in ASL. In: Meier, Richard P.; Cormier, Kearsy; and Quinto-Pozos, David (eds), pp. 88-111

Coulter, G.R. (1990) Emphatic Stress in ASL. In Fisher, Susan and Patricia Siple (eds.), 37-66

Dively, V.; Metzger, M.; Tabú, S. & Baer, A. (eds) (2000). Signed Languages. Discoveries from International Research, Washington DC: Gallaudet University Press.

Friedman, L.A., (1977) Formational Properties of American Sign Language, in On the other hand: new perspectives on American Sign Language. New York Academic press. 13-56

Goldsmith, J. (1992) A handbook of phonological theory. Oxford: Basil Blackwell

Fisher, S. & Siple, P. (1990) Theoretical Issues in Sign Language Research. Chicago, vol. I. The University of Chicago Press.

Herrero, A. & Alfaro, J.J. (1999) Fonología y escritura de la lengua de signos española. Estudios de Lingüística 13: 89-116

Herrero, A. (2003) Escritura alfabética de la lengua de signos española. Alicante, Publicaciones de la Universidad de Alicante.

Jakobson, R. (1988) Obras selectas, tomo I. Madrid, Gredos

Kanisza, G. (1980) Gramatica del vedere. Saggi su percezione e gestalt. Bologne, Il Mulino. Ed. esp. Barcelona, Paidós, 1986

Liddell, S.K. (1984) THINK and BELIEVE: Sequentiality in American Sign Language signs. Language 60: 372-390

Liddell, S.K. (1989) American Sign Language: The Phonological base. Sign Language Studies 64: 195-278

Liddell, S.K. (1990a) Four functions of a Locus: Re-examining the structure of space in ASL. In Lucas, C. (ed) Sign Language Research: Theoretical Issues.Washington, Gallaudet University Press, 176-198

Liddell, S.K. (1990b) Structures for Representing Handshape and local Movement at the Phonemic Level. In Fisher, Susan and Patricia Siple (eds.), 37-66

Malmberg, B. (1955) The phonetic basis for syllable division. Studia Linguistica 9: 80-87

Meier, R.P.; Cormier, K. & Quinto-Pozos, D. (eds) (2002), Modality and structure in signed and spoken languages. Cambridge: Cambridge University Press,

Olson, D. & Torrance, N. (eds.) (1991) Literacy and orality. Cambridge, Cambridge University Press. Ed. esp. Barcelona, Gedisa, 1998

Perlmutter, D.M. (1992) Sonority and syllable structure in American Sign Language. Linguistic Inquiry 23: 407-442

Pinedo, F. J. (1989) Nuevo diccionario gestual español. Madrid, C.N.S.E.

Sampson, G. (1985) Writing Systems. London: Century Hutchinson Limited. Ed. esp. Barcelona, Gedisa, 1997.

Sandler, W. (1986) The Spreading Hand Autosegment of American Sign Language. Sign Language Studies 50:1-28

Sandler, W. (1989) The Phonological Representation of the Sign. Dordrecht, Foris

Sandler, W. (1990) Temporal Aspects in ASL Phonology, in Fisher, S. & Siple, P. (eds.), 7-36

Stokoe, W.C. (1960) Sign language structure: An outline of the visual communication systems of the American Deaf. Burtonsville, Linstok Press, 1993

Straka, G (1963) La division des sons du langage en voyelles et consonnes peut-elle être justifiéé? Travaux de Linguistique et Littérature, I, 17-74

Wilbur, R.B. (1990) Why Syllables? What the Notion Means for ASL Research. In Fisher, S. & Siple, P. (eds.), 81-108

Wilbur, R.B. (1999) Stress in ASL: Empirical evidence and linguistic issues. Language and Speech 42: 229-250

Synthesis of Virtual Reality Animations from SWML using MPEG-4 Body Animation Parameters

Maria Papadogiorgaki[*], Nikos Grammalidis[*], Nikos Sarris[†] and Michael G. Strintzis[*]

[*]Informatics and Telematics Institute

1st Km Thrermi-Panorama Road

57001 Thermi-Thessaloniki, Greece

E-mail: {mpapad, ngramm, strintzi}@iti.gr

[†]Olympic Games Organizing Committee Athens 2004

Iolkou 8 & Filikis Etairias

14234 Nea Ionia, Greece

E-mail: NESarris@athens2004.com

Abstract

This paper presents a novel approach for generating VRML animation sequences from Sign Language notation, based on MPEG-4 Body Animation. Sign Language notation, in the well-known *SignWriting* system, is provided as input and is initially converted to *SWML (SignWriting Markup Language),* an XML-based format which has recently been developed for the storage, indexing and processing of SignWriting notation. Each *sign box* (basic sign) is then converted to a sequence of *Body Animation Parameters (BAPs)* of the *MPEG-4 standard*, corresponding to the represented gesture. These sequences, which can also be coded and/or reproduced by MPEG-4 BAP players, are then used to animate H-anim compliant VRML avatars, reproducing the exact gestures represented in sign language notation. Envisaged applications include producing signing avatars for interactive information systems (Web, E-mail, info–kiosks) and TV newscasts for persons with hearing disabilities.

1. Introduction

The *SignWriting* system is a writing system for deaf sign languages developed by Valerie Sutton for the Center of Sutton Movement Writing, in 1974 [1]. A basic design concept for this system was to represent movements as they are visually perceived, and not for the eventual meaning that these movements convey. In contrast, most of the other systems that have been proposed for writing deaf sign languages, such as *HamNoSys (the Hamburg Notation System)* or the *Stokoe system* employ alphanumeric characters, which represent the linguistic aspects of signs. Almost all international sign languages, including the American Sign Language (ASL) and the Brazilian Sign Language (LIBRAS), can be represented in the SignWriting system. Each *sign-box* (basic sign) consists of a set of graphical and schematic symbols that are highly intuitive (e.g. denoting specific head, hand or body postures, movements or even facial expressions). The rules for combining symbols are also simple, thus this system provides a simple and effective way for common people with hearing disabilities that have no special training in sign language linguistics, to write in sign languages. Examples of SignWriting symbols are illustrated in Figure 1.

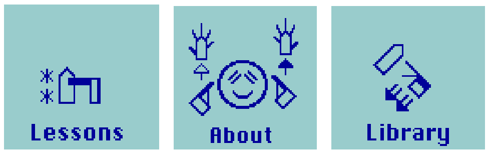


Figure 1: Three examples of representations of American Sign Language in SignWriting system.

An efficient representation of these graphical symbols in a computer system should facilitate tasks as storage, processing and even indexing of sign language notation. For this purpose, the SignWriting Markup Language (SWML), an XML-based format, has recently been proposed [7]. An online converter is currently available, allowing the conversion of sign-boxes in SignWriting format (produced by SignWriter, a popular SignWriting editor) to SWML format.

Another important problem, which is the main focus of this paper, is the visualization of the actual gestures and body movements that correspond to the sign language notation. A thorough review of state-of-the art techniques for performing synthetic animation of deaf signing gestures has been presented in [5]. Traditionally, dictionaries of sign language notation contain videos (or images) describing each sign-box, however the production of these videos is a tedious procedure and has significant storage requirements. On the other hand, recent developments in computer graphics and virtual reality, such as the new Humanoid Animation (H-Anim) [9] and MPEG-4 SNHC [3] standards, allow the fast conversion of sign language notation to Virtual Reality animation sequences, which can be easily visualized using any VRML-enabled Web browser.

In this paper, we present the design, implementation details and preliminary results of a system for performing such a visualization of sign-boxes, available in SWML. The proposed technique first converts all individual symbols found in each sign box to a sequence of MPEG-4 Body animation parameters. The resulting sequences can be used to animate any H-anim-compliant avatar using an MPEG-4 SNHC BAP player provided by EPFL [4]. The system is able to convert all hand symbols as well as the associated movement, contact and movement dynamics symbols contained in any ASL sign-box. Although only manual (hand) gestures are currently supported, we plan to implement other body movements (e.g. torso) as well as facial animation in the near future. The proposed technique has significant advantages:

- Web- (and Internet-) friendly visualization of signs. No special software has to be installed,
- Allows almost real-time visualization of sign language notation, thus enabling interactive applications,
- Avatars can easily be included in any virtual environment created using VRML, which is useful for a number of envisaged applications, such as TV newscasts, automatic translation systems for the deaf, etc.
- Efficient storage and communication of animation sequences, using MPEG-4 coding techniques for BAP sequences.

Significant similar work for producing VRML animations from signs represented in the HamNoSys transcription system to VRML has been carried out by the EC IST ViSiCAST project [6], and its follow-up project "E-Sign"[10]. Current extensions of HamNoSys are able to transcribe all possible body postures, movements and facial expressions [11] and significant work towards supporting MPEG-4 BAPs has been made. The main contribution of the proposed approach in this paper is the attempt to work towards the same direction for the most common and popular representation of Sign Languages, which is the SignWriting notation system.

The paper is organized as follows: Section 2 provides an introduction to SWML and describes how our application extracts information from SWML files. In Section 3, the proposed technique for converting sign boxes to MPEG-4 Body Animation Parameters is described. The synthesis of animations for H-anim avatars is outlined in Section 4, while discussion and future work is presented in Section 5.

2. Introduction to SWML and parsing of SWML files

SWML [2] is an XML-based format described by the SWML DTD (currently version 1.0 draft 2)[7]. The DTD specifies two types of SWML documents: *sw_text* (sign language text generated e.g. an SWML editor or converter) and *sw_table* (sign language database or dictionary generated by an SWML aware application).

- An sw_text document consists of *sign_boxes* and *text_boxes*, where each sign box consists of a set of *symbols* and each text box contains an alphanumeric string.
- An sw_table document consists of table of entries, where each entry consists of a *sign_box* and a corresponding *gloss* (a sequence of fields containing descriptions for this sign box in an oral language).

Each symbol is specified in SWML using the following fields:

a) A *shape number* (integer) specifying the shape of the symbol,
b) A *variation* parameter (0 or 1 for hand symbols / 1,2 or 3 for movement and punctuation symbols) specifying possible variations (complementary transformations) of the symbol,
c) A *fill* parameter (0,1,2 or 3 for hand and punctuation symbols / 0,1 or 2 for movement symbols) specifying the way the shape is filled, generally indicating its facing to the signer,
d) A *rotation* parameter (0-7) specifying a counter-clockwise rotation applied to symbol, in steps of 45 degrees,
e) A *transformation flip* parameter (0 or 1) indicating whether the symbol is vertically mirrored or not, relatively to the basic symbol and, finally,
f) The *x* and *y* coordinates of the symbol within the sign box.

For sign synthesis, the input for the sign synthesis system consists of the SWML entries of the sign boxes to be visualized. For each sign box, the associated information corresponding to its symbols is parsed. Information related to symbols that are supported by the sign synthesis application, i.e. hand symbols as well as corresponding movement, contact and movement dynamics symbols, is then used to calculate the MPEG-4 Body Animation Parameters.

3. Conversion of Sign Boxes to MPEG-4 Body Animation Parameters

The issue of body modeling and animation has been addressed by the Synthetic/Natural Hybrid Coding (SNHC) subgroup of the MPEG-4 standardization group [3]. More specifically, 168 Body Animation Parameters (BAPs) are defined by MPEG-4 SNHC to describe almost any possible body posture. Most BAPs denote angles of rotation around body joints. In this section, the proposed system to convert symbols contained in a SWML sign box to BAP sequences will be presented.

Currently, symbols from the 1995 version of the Sign Symbol Sequence (SSS-1995) are supported. This sequence comprises an "alphabet" of the SignWriting notation system, while true images (in gif format) of each symbol contained in this sequence are available in [2]. The proposed system is able to convert

- All 106 hand symbols,

- All 95 (hand) movement symbols and
- Two punctuation symbols (180,181), which contain synchronization information.

Other punctuation symbols as well as symbols that represent face expressions and face, torso and shoulder movements (43 symbols) are currently ignored (not decoded) by the system.

The conversion starts by first examining the symbols contained within the input sign box. If no symbols describing dynamic information such as hand movements, contact or synchronization exist, the resulting BAP sequence corresponds to just one frame (i.e. a *static gesture* is reproduced). Information provided by the fields of the (one or two) hand symbols, contained in the sign box, is used to specify the BAPs of the shoulder, arm, wrist and finger joints. On the other hand, if symbols describing dynamic information exist, the resulting BAP sequence contains multiple frames, describing animation key-frames (i.e. a *dynamic gesture* is reproduced). The first key-frame is generated by decoding the existing hand symbols, as in the case of static gestures. Since the frame rate is constant and explicitly specified within a BAP file, the number of resulting frames may vary, depending on the complexity of the described movement and its dynamics. Synchronization symbols and contact also affect the represented movement and in some cases require special treatment.

Smooth and natural-looking transitions from and between the neutral body position and the body position corresponding to a static gesture (or the start and end frames of a dynamic gesture) is achieved by generating additional intermediate frames using a "hierarchical" BAP interpolation procedure: intermediate BAP sets (frames) are generated to consecutively move first the arms, then the wrist and finally the fingers from their previous positions to their new positions.

A block diagram of the proposed system is illustrated in Figure 2, while additional details about the generation of BAPs for static and dynamic gestures are provided in the following Subsections.

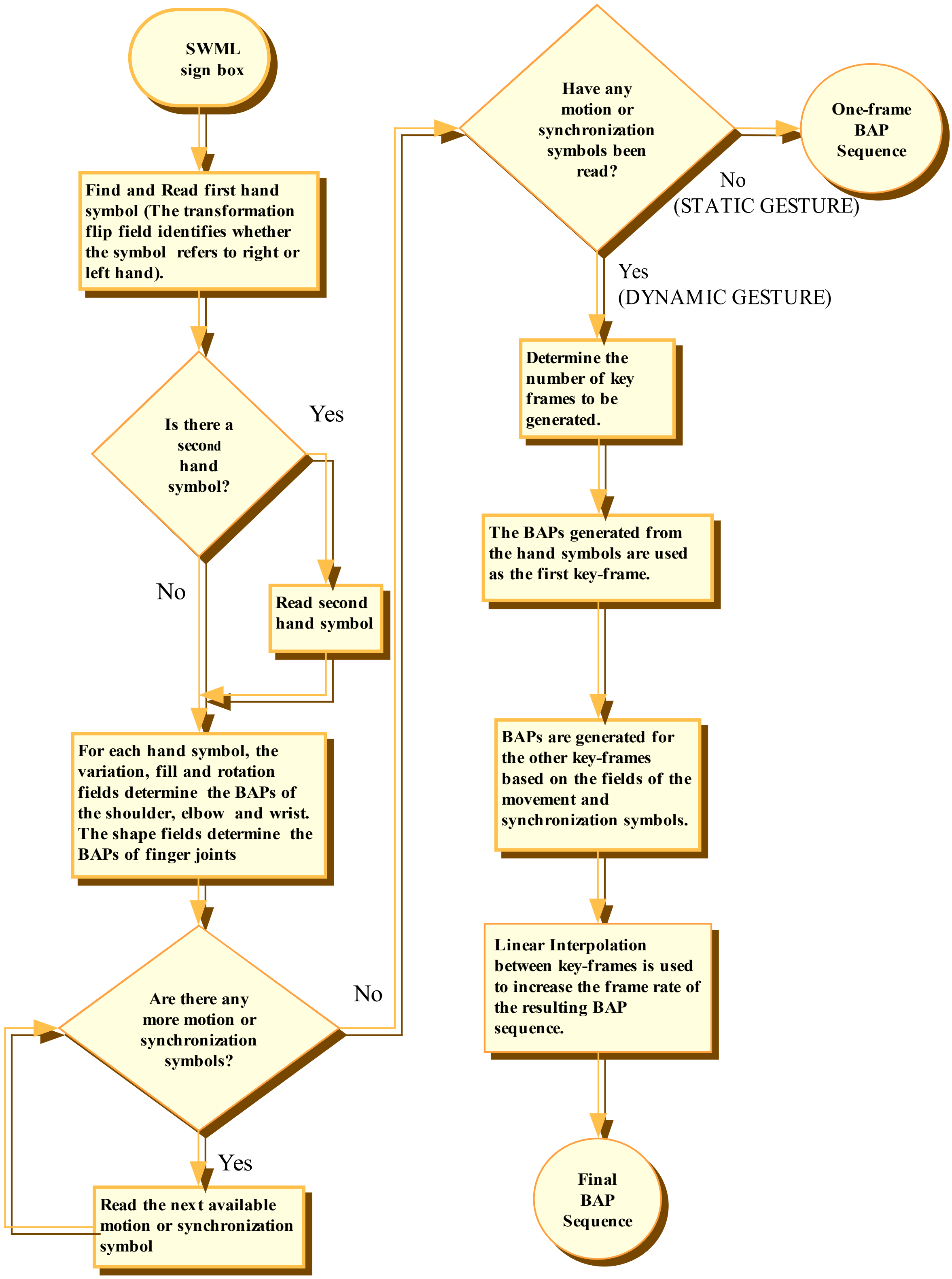


Figure 2: A block diagram of the proposed system.

3.1. Static gestures

The SignWriting system allows various transformations to be applied to a basic symbol. A hand symbol for example can exist in many different postures with bent fingers etc, represented with different shape numbers. Also the signer may either see his palm, the back of his palm or the side of his palm (Figure 3).

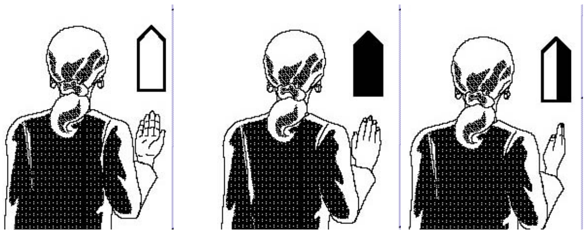


Figure 3: The signer sees a) his palm, b) the back of his palm c) the side of his palm.

As seen in Figure 4, the hand may either be parallel with the wall (wall plane) or with the floor (floor plane).

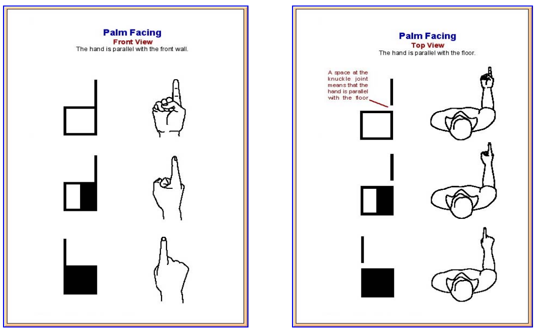


Figure 4: a) Hand is parallel with the wall plane b) Hand parallel is with the floor plane

The position of the palm may also change due to a rotation around the wrist joint. Furthermore, a "flipped" symbol represents a symbol that is "mirrored" around the vertical axis. This means that it actually describes a posture of the other hand. A hand symbol and its flipped version are illustrated in Figure 5.

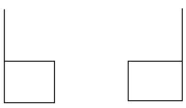


Figure 5: A basic hand symbol and its flipped version.

In the following, the procedure to extract useful information from the SWML representation of a hand symbol is summarized:

Initially, the binary "transformation flip" parameter is used to identify whether the symbol corresponds to the left or right hand. Then the fill and variation parameters of each symbol are used to determine the animation parameters of the shoulder and elbow joints:

- If (variation,fill)=(0,0),(0,1) or (1,3) then the axis of the arm is parallel to the floor (floor plane).
- If (variation,fill)=(1,0),(1,1) or (1,2) then the axis of the arm is parallel to the human body (wall plane)
- If (variation,fill)=(1,0) or (1,3) then the signer sees his palm
- If (variation,fill)=(1,1) or (0,0) then the signer sees the side of his palm
- If (variation,fill)=(1,2) or (0,1) then the signer sees the back of his palm

In addition, the rotation parameter is used to determine the animation parameters of the wrist joint:

- If the signer sees the side of his palm, the rotation value (multiplied by 45 degrees) is used to define the **R_WRIST_FLEXION BAP** (for the right hand) or the **L_WRIST_FLEXION BAP** (for the left hand).
- In the other two cases (signer sees his palm or the back of his palm), the rotation value (multiplied by 45 degrees) is used to define the **R_WRIST_PIVOT BAP** (for the right hand) or the **L_WRIST_PIVOT BAP** (for the left hand).

Finally, the symbol shape number is used to specify the animation parameters corresponding to finger joints, using look-up tables of BAP values corresponding to each symbol.

If the sign box contains a second hand symbol, similar procedures are used to extract the body animation parameters of the other hand. After the processing of all existing hand symbols, all body animation parameters corresponding to shoulder, elbow, wrist and finger joints are determined and stored.

3.2. Dynamic gestures

A movement symbol may exist in many forms describing either simple or complex movements. Movement can be either parallel to the wall plane or to the floor plane. Furthermore, as can be seen in Figure 6a, movement symbols for the left and right hand have different representations. When the movement is associated with the right (left) hand, the arrow representing its direction has a dark (light) arrowhead. When both hands are simultaneously moving to the same direction as a group, the representation of the movement is done using a neutral arrowhead, which is neither dark nor light. In some cases, the size of a movement symbol is used to specify the duration (i.e. the speed) of the hand.

For example, the arrow symbol in Figure 6b is illustrated in three different sizes: the first represents a fast movement forward, the second represents a movement forward with normal speed and the last represents a slow movement forward.

dynamic gesture are then generated from a specific set of functions.

Figure 6: Three versions of a symbol specifying:

a) movements of different hands, b) movements

with different time durations.

MPEG-4 standard allows the description of human body movement using a specific set of body animation parameters corresponding to each time instant. Systems like SignWriting that use a high level animation description define movement by specifying a starting and an ending position, in case of simple motion with constant velocity, or the full trajectory, in case of more complex motion. However, the description of complex motion is also possible by specifying a number of intermediate key-frames. In the following, the procedures for generating these BAP key-frames are briefly described.

3.2.1. Generation of BAP key-frames

When all movement description symbols have been identified, the shape number field identifies their shapes (i.e. the type of movement). First, the total number of key-frames to be produced is specified, based on the number and nature of the available movement, movement dynamics, contact, and synchronization symbols. More specifically, a look-up table is used to define an initial number k of key frames for each movement symbol. Furthermore, the fill parameter specifies whether the motion is slow, normal or fast. In addition, some symbols explicitly specify the movement duration. For this reason, a classification of such symbols into three categories has been defined and a different duration value D is defined for each category:

- Slow motion (D=3)
- Normal motion (D=2)
- Fast motion (D=1)

The total number of frames to be generated when only one motion symbol exists is $N=kDP$, where P is a fixed multiplier (e.g. P=10). If the number of such symbols is more than one, the total number of key-frames is the maximum between the numbers of key-frames, corresponding to each symbol. Finally, if the sign box contains a contact symbol, the total number of frames is increased by two (in case of simple contact) or four (in case of double contact).

The initial key-frame is generated by decoding the available hand symbols, exactly as in the case of static gestures. The rotation and transformation flip fields specify the exact direction of movement. Also, the variation field specifies whether the right or the left hand performs the movement. Using information from all available movement, contact and synchronization symbols, the other BAP key-frames of the specific dynamic gesture are then generated from a specific set of functions.

3.2.2. BAP Interpolation

Finally, when the BAPs for all key-frames have been computed, BAP interpolation is used to increase the frame rate of the resulting BAP sequence. This interpolation procedure results to smoother transitions between key frames.

Interpolation is generally achieved by approximating the motion equation using a mathematical function and then re-sampling this function to obtain the desired intermediate positions at intermediate time instants. Various interpolation functions can be selected in order to improve results. Since body animation parameters represent rotations around specific joints, quaternion interpolation was seen to provide good results [8], but the complexity of the method is increased. For this reason, a linear interpolation technique was applied, which was seen to be very efficient for most signs, since key-frames have been selected so as to simplify the movement description between consecutive key-frames.

3.2.3. Synchronization (Movement Dynamics) Symbols: A special case

The sign box may also contain one of the three supported synchronization (movement dynamics) symbols (180,181 and 182). These symbols as well as their fields and interpretation are described below:

Shape number=180

- Variation=1, fill=0: simultaneous line (both hands move at the same time)
- Variation=1, fill=1:alternating lines (the right hand move in one direction while the left move simultaneously in the opposite direction)
- Variation=1, fill=2: un-even alternating (one hand moves while the other is still then the second hand moves while the first remains still)
- Variation=1, fill=3, rotation=0: slow movement
- Variation=1, fill=3, rotation=4: smooth movement

Shape number=181

- Variation=1, fill=0: tense movement
- Variation=1, fill=1: tense movement with emphasis
- Variation=1, fill=2: relaxed movement
- Variation=1, fill=3: relaxed movement with emphasis

Shape number=182

- Variation=1, fill=0: fast movement
- Variation=1, fill=1: fast movement with emphasis

These synchronization symbols are handled in a similar way as movement symbols but an exception exists for the “Un-even alternating” symbol, where first one hand moves, while the other hand is still and then the opposite. To handle this case the total number of key frames is doubled ($N=2kDP$). To produce the first kDP frames, BAPs are generated only for the first hand, so the second hand remains still. In the following, BAPs are generated only for the second hand, to produce the next *kDP* frames, so the first hand remains still.

4. Synthesis of animations using h-anim avatars

The "EPFLBody" BAP player [4], developed by the École Polytechnique Fédérale Lausanne (EPFL) for the Synthetic and Natural Hybrid Coding (SNHC) subgroup of MPEG-4 was used to animate H-anim-compliant avatars using the generated BAP sequences. Since most BAPs represent rotations of body parts around specific body joints, this software calculates and outputs these rotation parameters as animation key-frames to produce a VRML ("animation description") file that can be used for animating any H-anim-compliant VRML avatar. Two frames from resulting animations are illustrated in Figure 7

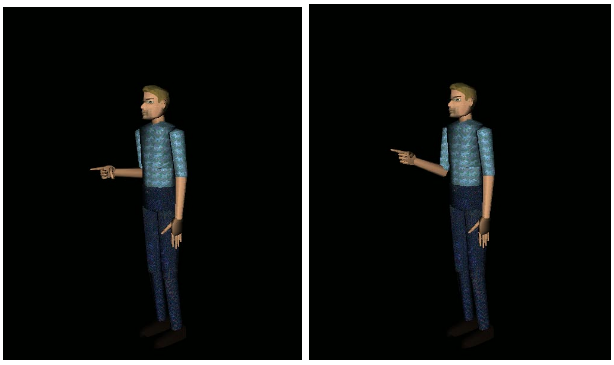


Figure 7: Animation of the "You" sign in ASL using an H-anim avatar

By including a VRML TouchSensor Node within the VRML file describing the H-anim avatar, the viewer can interactively start and/or replay the animation sequence, by clicking on the avatar. The viewer can also interact by zooming in and out to any specific body region and/or by rotating and translating the model within the 3-D space, in order to fully understand the represented sign.

Furthermore, further evaluation of the proposed sign synthesis system was possible by developing an online system [12] for converting text to Sign Language notation and corresponding VRML animation sequences for H-anim compliant avatars. The application, whose interface is illustrated in Figure 8, is currently based on a 3200-word SWML dictionary file, obtained by the SWML site [2], which has been parsed and inserted into a relational database. The user is allowed to enter one or more words, which are looked up in this dictionary. If more than one entry is found, all possible interpretations are presented to the user, so that he can choose the desired one. On the other hand, if no entries are found for a specific word, the word is decomposed using its letters (finger-spelling). In any case, the user may choose whether to include a particular term to the selected terms to be used for sign synthesis or not. The user then selects an H-anim compliant avatar, which is used for sign synthesis of the selected term or terms. Furthermore, the user may produce and display the corresponding sign(s) in SignWriting format (in PNG format) and SWML for a specific term or the selected terms.

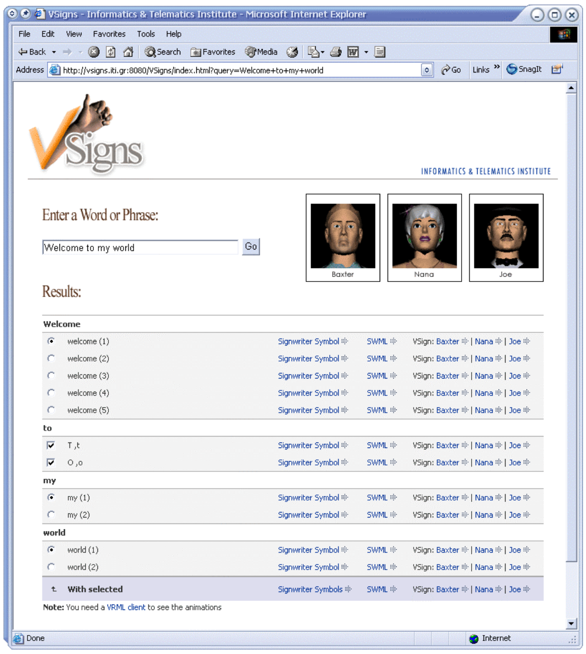


Figure 8: Example query: "Welcome to my world". The user may then select the desired terms and then produce and display sign synthesis results using the selected words or the entire phrase, using any of the available H-anim avatars.

This experimental Web application has already allowed us to identify significant problems with the synthesis of static and dynamic gestures, which have to be solved in the future, e.g. when contacts and complex movements are involved. A major problem that has to be solved occurs when the sign-box contains contact symbols. In that case the touch between the hands, or the hand and the face is difficult to be achieved. Problems may also occur for complex movements, when the inclinations of the hand joints, which have been estimated in each key frame, are not accurate enough for the exact description of the movement. Both problems can be solved in the future by using inverse kinematics methods.

Further evaluation is planned for the future, using Greek and International SignWriting users, and attempts will be made to solve the problems that have been observed or will be observed in the future. Although these problems indicate that much more work is needed for correct synthesis of all signs, we believe that with this Web tool, a very important step towards automatic Text to Sign synthesis has been made.

5. Discussion and Future work

A novel approach for generating VRML animation sequences from Sign Language notation, based on MPEG-4 Body Animation has been presented. The system is able to convert almost all hand symbols as well as the associated movement, contact and movement dynamics symbols contained in any ASL sign-box.

As stated in the introduction, we plan to support non-manual body movements as well as facial animation within the near future. Facial animation will be represented by MPEG-4 Facial Animation Parameters, while animation of H-anim compliant avatars using simultaneous face and body animation has been already successfully implemented. A problem with using Facial Animation Parameters is that most of them, in contrast to BAPs, describe complex non-rigid motions, and therefore most existing FAP player implementations are model-dependent. Furthermore, the resulting VRML animations are more complicated since they contain numerous CoordinateInterpolator nodes (one per face model vertex). Therefore, the computational demands for the hardware that is reproducing these animations are increased.

Finally, a short-term goal is to design practical applications of the proposed system, either as a "plug-in" to existing applications (e.g. sign language dictionaries) or as a stand-alone tool for creating animations for TV newscacts (e.g. weather reports). Particular emphasis will be given in applications that can be used and evaluated by the Greek Sign Language community, thus a dictionary of Greek Sign language, in SignWriter notation, is planned to be supported in the near future.

6. Acknowledgement

This work was supported by FP6 IST Network of Excellence "SIMILAR" -"The European taskforce creating human-machine interfaces SIMILAR to human-human communication". The authors would also like to thank Lambros Makris for designing and developing the "Vsigns" Web Page.

References

[1] Official SignWriting site, http://www.signwriting.org/

[2] Official Site of SWML, http://swml.ucpel.tche.br/

[3] Moving Pictures Experts Group, (1999). Generic coding of audio-visual objects - Part 2: Visual, MPEG Document ISO/IEC JTC1/SC29/WG11 N3056. Maui.

[4] Fabrice Vergnenegre, Tolga K. Capin, and D. Thalmann (1999). Collaborative virtual environments-contributions to MPEG-4 SNHC. ISO/IEC JTC1/SC29/WG11 N2802, http://coven.lancs.ac.uk/mpeg4/

[5] A B. Grieve-Smith (2001). SignSynth: A Sign Language Synthesis Application Using Web3D and Perl. Gesture Workshop, London, pp. 134-145.

[6] R. Kennaway (2001). Synthetic Animation of Deaf Signing Gestures. Gesture Workshop, pp. 146-157, London.

[7] Antonio Carlos da Rocha Costa, Cracaliz Pereira Dimuro (2001). Supporting Deaf Sign Languages in Written Form on the Web. The SignWriting Journal, Number 0, Article 1, July. http://gmc.ucpel.tche.br:8081/sw-journal/number0/article1/index.htm.

[8] M. Preda, F. Preteux (2001). Advanced virtual humanoid animation framework based on the MPEG-4 SNHC standard. Proceedings EUROIMAGE International Conference on Augmented, Virtual Environments and Three-Dimensional Imaging (ICAV3D'01), Mykonos, Greece, pp. 311-314.

[9] Humanoid Animation Standard Group. Specification for a Standard Humanoid: H-Anim 1.1. http://h-anim.org/Specifications/H-Anim1.1/

[10] Official site of E-sign (Essential Sign Language Information on Government Networks) project. http://www.visicast.sys.uea.ac.uk/eSIGN/

[11] Th. Hanke (2002). iLex - A tool for sign language lexicography and corpus analysis. In: Proceedings of the Third International Conference on Language Resources and Evaluation, Las Palmas de Gran Canaria, Spain., pp. 923–926.

[12] Vsigns page, http://vsigns.iti.gr

Multipurpose Design and Creation of GSL Dictionaries

Eleni Efthimiou, Anna Vacalopoulou, Stavroula-Evita Fotinea, Gregory Steinhauer

ILSP-Institute for Language and Speech Processing
Artemidos 6 & Epidavrou, GR 151 25, Maroussi, Greece
{eleni_e, avacalop, evita, stein}@ilsp.gr

Abstract

In this paper we present the methodology of data collection and implementation of databases with the purpose to create extensive lexical and terminological resources for the Greek Sign Language (GSL). The focus is on issues of linguistic content validation, multipurpose design and reusability of resources, exemplified by the multimedia dictionary products of the projects NOEMA (1999-2001) and PROKLISI (2002-2004). As far as data collection methodology, DB design and resources development are concerned, a clear distinction is made between general language lexical items and terms, since the creation of resources for the two types of data undergoes different methodological principles, lexeme formation and usage conditions. There is also reference to content and interface evaluation mechanisms, as well as to basic linguistic research carried out for the support of lexicographical work.

1. Introduction

A basic requirement for the treatment of signs or sign streams as linguistic input for NLP and for the development of applications that make use of linguistic data, is the existence of adequate linguistic resources in the form of electronic lexical databases and computational grammars.

The Greek Sign Language (GSL) has only recently started to be subject to systematic linguistic analysis. This is, on one hand, due to the fact that it was not until 2000 (Act 2817) that GSL was recognized by the Greek Parliament as an official language of the Greek State. On the other hand, this interest is directly connected to the development of technologies, which enabled the creation of electronic linguistic resources (including lexicons, grammars and sign language corpora) for languages that are uttered in the three-dimensional space (see also Efthimiou et al., 2004). Such resources can nowadays be adequately stored, retrieved and represented, exploiting the ability of current systems to incorporate various multimedia functionalities for the generation of signs, into a single platform.

2. GSL lexicography: the background

In contrast to other sign language systems, i.e. the ASL (Tennant & Gluszak Brown, 1998 ; Wilcox et al., 1998), systematic lexicographical work in respect to GSL has started only recently, within the framework of the NOEMA project (1999-2001).

This was the first attempt to create multipurpose reusable linguistic resources for GSL. Part of the project description was the creation of a digital sign stream narration corpus and an electronic dictionary of basic GSL vocabulary. The spin-off products of that project, among which are a 3,000 entry multimedia bilingual dictionary (GSL-Greek) of basic vocabulary and a multimedia children's dictionary of GSL (Kourbetis & Efthimiou, 2003), reflect the methodology for creating linguistic resources followed, the content and interface evaluation mechanism adopted, as well as the basic linguistic research carried out to support the lexicographical work (NOEMA Project, 2001).

The knowledge acquired with respect to the morpho-phonological operations the formation of simple and complex signs allowed for: a) the construction of rules for creating new valid signs, b) the denomination of relevant terms and c) the classification of GSL linguistic resources into terminological lists. All these have significant impact on the development of both communication and educational tools using technologies which allow the 3D representation of linguistic content.

3. Methodological principles of vocabulary formation

The initial steps of our work on GSL vocabulary included a survey of the existing lexicography (Logiadis & Logiadis, 1985) and syntax literature. It came out that the available knowledge of GSL was only based on individual fragmentary attempts. These usually lacked scientific criteria, did not derive from systematic scientific analysis and generally involved the creation of some kind of lexicon. This fact is directly connected with the prevailing assumption that GSL is not an autonomous linguistic system but, rather, some kind of representation of aural Greek.

Consequently, the creation of lexical resources had to take into serious consideration the linguistic material that would serve as the basis for the lexicographical work (Johnston & Schembri, 1999) and which should reflect linguistic synchrony, also allowing for an adequate grammatical description of GSL (Bellugi & Fischer, 1972).

Next, we will present the methodologies adopted for compiling two vocabulary lists: a general purpose basic vocabulary of 3,000 initial signs and a vocabulary of basic computer-skills terminology.

In both cases, extensibility and reusability were the main design principles, whereas lack of previous linguistic resources dictated specific methodological approaches to data collection (for the general purpose vocabulary), as well as to new sign formation (for the computer-skills terminology list).

3.1. Methodology of creation of a general purpose basic vocabulary: data collection

The first step of this task mainly involves the compilation of the basic sign vocabulary[1] of GSL. In the process of compiling a list of 3,000 basic signs of GSL without an appropriate corpus available, a decision had to be made as to whether statistical frequencies, every day use or vocabulary lists taught to young children would constitute our data.

In order to overcome the lack of GSL resources, we comparatively studied the proposed basic vocabularies or 'defining vocabularies' of three well analyzed aural languages: English, French and German (Mueller et al., 1995 ; Gougenheim, 1958 ; Longman Dictionary of Contemporary English). Based on this study, we gathered a core 3,650 lemma list, which was, then, compared to two other lists:

- the first one, containing 1,850 words, was provided by the Hellenic Federation of the Deaf (HFD) and derived from a previously videotaped and lemmatized corpus to serve as basic study material for GSL;
- the second one contained the 2,100 most frequent words in the Hellenic National Corpus (HNC), an electronic corpus of general Greek developed by ILSP, which contained 13,000,000 words at the period of study.

The HNC (1999) word list is of significant importance, given that it contains words corresponding to existing appearances in text corpora. On the other hand, the words that consist the basic vocabularies of different languages carry an even heavier weight because they allow reference to a set of concepts rather than isolated words. Such concepts may be viewed as basic in respect to everyday communication. Since we proposed a concept-based approach to vocabulary building, we had to take into account the issue of the representation of these concepts through different grammatical categories. We noticed that in the vocabulary lists included in our study, concepts were represented either by a single or by more than one grammatical category, without following a systematic way of listing (i.e. in one case, the proposed representation involves *basic/base*(v) vs. *base*(n)/*base*(v) and in another *difference/differ* vs. *difference/different*).

In the case of GSL vocabulary, we either adopted the words suggested by HFD or followed suggestions made by individual native GSL informants. Specific grammatical categories were further excluded from the GSL list on the basis of the numerical restriction of 3,000 signs. Subject to this exclusion were adverbs (unless no equivalent adjective was available) and passive verb forms and participles (unless the latter had an adjectival function in the language).

As a result, a 2,800 concept list was formed, which was then presented to HFD for comments, enrichment with concepts specific to deaf communication and video recording (Efthimiou & Katsoyannou, 2001). For every concept on the proposed list three parameters are true:

- they have a high frequency rate in the vocabulary of Greek according to HNC data;
- they are included in at least two of the proposed basic vocabularies we took into account (Figure 1);
- they can be expressed by words of more than one grammatical category (i.e. love(n)/love(v)) or by a concatenation of synonyms (i.e. angry-furious).

The aim of this procedure was to form the basic sign list of GSL as used by native signers without being biased by external parameters. For this reason, our informants were asked to propose synonym or antonym signs for concepts, wherever possible, so that semantic relations be stated by means of GSL mechanisms rather than via influence from spoken Greek or other language systems.

3.2. Methodology of development of terminological resources

As far as GSL terminological resources design is concerned, we had to take into account that the introduction of specific concept systems in the language means creating new term systems for representing these concepts (Sager, 1994 ; Otman, 1996). In the initial stage of defining the methodology for term formation, we focused on the principle that new term denominations, term signs in our case, should incorporate and demonstrate the following properties innate to the language (Gee & Goodhart, 1985):

- GSL mechanisms of vocabulary organization;
- GSL mechanisms of word formation;
- GSL lexical unit function in sign stream production.

The task of term list formation (Rey, 1995) incorporates, to a considerable extend, the characteristics and conditions of lexicographical work. However, there is a crucial point of differentiation, as the items included in a terminology list carry a given semantic value only within a specific context of use, outside which they may carry different meaning or no meaning at all.

Furthermore, terms are one-to-one representations of concepts, which are organized into systems (Rey, 1996) and, in contrast to other lexical items, may consist of complex syntactic and/or semantic units which are formed not merely by linguistic but also by other (i.e. mathematical) symbols or a combination of them (Wright & Strehlow, 1995).

The primary task in terms of the initial linguistic data collection was defining the field of coverage (Sager, 1990). This was followed by a study of the content of term intensive corpora on the selected fields of knowledge. The result was the extraction of a set of concepts for each field. Our example case is the field of computer-skills terminology. In this specific case, the language of initial knowledge creation is English. As a result, a considerable proportion of the terms, denominating the relevant concepts, are transferred either directly or indirectly from English into receiver languages, such as Greek. Consequently, the concept list of computer-skills terminology had, in our case, two existing representation equivalents in the context of spoken languages: a set of English terms (source language) and a set of their Greek translations (receiver language).

[1] One should notice that the notion of basic vocabulary is not uniformly defined in the relevant literature, which raises the issue of selecting the appropriate methodological approach to deal with the data.

The task was to create terms in GSL for the same concepts, allowing for the least possible influence by previously existing representations, while creating terminological items according to sign language word formation principles. This was a crucial prerequisite for the proposed denominations to be recognized by native signers as components of GSL with acceptable internal structure and specific cognitive content.
This task of concept denomination for the formation of a terminology list in GSL was undertaken by a working group of terminologists, computational linguists, computer scientists, GSL specialists and computer skills teachers which included members of the Greek Deaf Community.
The output of this group work was a list of video recorded terms, which were entered into a DB along with their Greek and English equivalents.

4. Organization of vocabulary databases

The internal organization of the lexical resources database differs from the one designed for storing terminological items with respect to lemma-related information as far as the expected functionality of resources is concerned. Thus, synonyms and antonyms (Figure 2) are included only in the case of general vocabulary, whereas standard GSL phonological features such as handshapes are included as lemma related information in both DBs. For the same reasons, lemmas in the terminology DB are related not only to field but also to sub-area of use, in order to allow for greater precision and clear lemma interrelations.

4.1. Design and development of the general purpose vocabulary DB

Given the specific goal of creating exhaustive reusable vocabulary resources of GSL, the design of the general purpose vocabulary DB incorporated a number of properties which include fields for:

- video recorded signs,
- grammatical category of lemmas,
- synonyms,
- antonyms (Figure 3),
- interpretations,
- lemma classification by thematic category,
- lemma translation into Greek and
- HamNoSys annotation features of lemma phonology (Prillwitz et al., 1989).

The DB was then enriched with lexical content following the methodology for data collection described above.
Experience gained by lemma analysis of the selected video signs enabled a number of assumptions regarding the morphological structure and sign formation mechanisms of GSL (Efthimiou & Katsoyannou, 2002). This knowledge provided the grounds for introducing new signs as in the case of GSL terminology items.
The implementation of the DB has already proven that the above structure allows for a multi-dimensional use of the resources created. The reusability of the general GSL vocabulary resources has already been tested by the fact that these resources provided the lexicographical content for a number of dictionary products. The same DB content also draws on on-going research with respect to efficient sign representation.

4.2. Design and development of the terminological DB

The design of the terminological resources DB is based on a term list, the formation of which was described in the methodology section 3.2 above. Each entry corresponds to a term and includes fields for:

- the video recorded term-sign,
- a video capture file serving as a visualized definition (Rousseau, 1983),
- the equivalent Greek term,
- the equivalent English term,
- a lemma identification code number,
- a code indicator corresponding to the basic handshape for the term-sign formation in GSL,
- a link to HamNoSys features other than the handshape, and
- sub-area fields in which each term is used.

In the case of computer-skills terminology, the sub-area fields include the following categories:

- General Notions,
- Word,
- Excel,
- Access,
- Internet Explorer,
- Power Point and
- Windows.

By adopting this architecture, the extensibility of the DB is guaranteed through the possibility of adding new terms, entry fields or terminology domains. Moreover, DB maintenance through addition, deletion or modification of term entries is possible without crucial or risky changes in terms of programming (Sowa, 2000).

5. Dictionary implementation

To exemplify the (re-)usability of the lexical resources discussed here, we make a short reference to two relevant products: a bi-directional (aural Greek-GSL and GSL-aural Greek) dictionary, compiled after a systematic survey of linguistic structure and a computer-skills trilingual dictionary (GSL-Greek-English).
As far as the dictionary making process is concerned, the organisation of entries was based upon the principle of usability in terms of the two user groups. Thus, each sign-lemma is followed by different defining / exemplification elements in both cases. In the general purpose dictionary (Efthimiou & Katsoyannou, 2001 ; 2002), entry structure provides the following set of information with respect to each GSL lemma:

- translation equivalent(s),
- an explanation in Greek,
- synonyms in GSL,
- antonyms in GSL,
- illustrative image (whenever possible),
- thematic category for lemma classification.

The inclusion of a Greek definition and translation helps non-native GSL signers enrich their vocabulary of modern Greek. At the same time, thematic categorization enables the learning of groups of signs which relate to each other semantically or pragmatically.
Lemma search is possible in the following manners:

- by order of handshapes within lemmas (Figure 4),
- by thematic category (e.g. «plant names»),
- by alphabetical order of the modern Greek translations.

Dictionary users perceive the special features of GSL in direct reference with Greek, while thematic fields function as a bridge between each sign and its Greek equivalent.

Concerning the terminology dictionary, as soon as the application starts, the items in the DB are processed so as to filter the lemmas corresponding to the user selection criteria (PROKLISI Project, 2003).

The lemma screen includes the following elements:

- thematic category,
- a list of every lemma in this category, from which users can select,
- the selected lemma in Greek,
- the selected lemma in English,
- a video-lemma in GSL,
- a list of all sub-area fields in which the selected lemma appears,
- a screen capture example of the term,
- a videotaped text in GSL with a concise presentation of the selected thematic category.

Users can access the content in the following ways:

- by the main handshape which forms the sign corresponding to each term. In this case, each sign is also accompanied by equivalents in both Greek and English, a list of thematic categories relevant to the term, a video presentation of the term, and a videotaped text with an introduction to the selected sub-area;
- by the Greek or English term equivalents in alphabetically ordered lists (Figure 5). The sign which corresponds to the selected term can appear either by clicking on the list or by typing it in, in one of the suggested languages. Items of information available for this search option include: a list of every sub-area in which the selected lemma appears, a video exemplifying the lemma and the videotaped text with an introduction to the selected thematic sub-area;
- by thematic sub-area. In this case, users can select among seven thematic categories (Figure 6) corresponding to the sub-areas in which computer-skills terminology is categorized. This option retrieves the corresponding terms in three lists of equivalents: GSL-Greek-English. Items of information available for this search option also include the other sub-areas in which the term appears, a video capture explanation of the term or an image, and an informative sign stream presentation of the selected sub-area.

6. Evaluation criteria and procedure

Evaluation procedures for both dictionary products were carried out by user groups of native GSL signers in real use environment. The basic vocabulary dictionary was tested in two rounds, in the context of various communicative situations. The evaluation body was composed of GSL native signers of various age groups, who were asked to use the dictionary in school, work and home environment and complete an evaluation criteria list. The related questionnaire contained 26 multiple choice questions and 5 free input slots. The main evaluation criteria comprised educational and communication needs, type of profession, source that disseminated the NOEMA product, interface design (screen organization, menus, help provided), efficiency of information accompanying the entry for each sign, adequacy of information introducing general aspects of GSL grammar incorporated in the product, period for getting used to navigating through the product and possible recommendations for future versions. The output of that first circle of evaluation served as feedback for making improvements to the final dictionary product. The second evaluation step followed the same methodology, with the purpose of verifying the acceptance of the product by the Greek Deaf Community. More information on the evaluation of the basic vocabulary dictionary can be found at the related project deliverable (NOEMA, 2001).

A first version of the computer-skills terminology dictionary was experimentally introduced as an education support tool in a continuous education class. Comments on both system functionality and content efficiency were incorporated in the final product version to be released on 30th March 2004.

7. Future research & development goals

Future development efforts in respect to both platforms (basic vocabulary dictionary and computer terminology dictionary) include investigation of the possibility of implementing smarter search options, in relation to the ongoing extension of the basic vocabulary DB content. Efficient sign-based user look-up features will also be incorporated along with fuzzy search capabilities (as proposed, for instance, by Wilcox et al. (1994)).

Based on the proposed methodology for the creation of the computer-skills terminology dictionary, other specialized dictionaries, intended to serve knowledge transfer in several areas of interest, are foreseen to be created, in order to meet a wider range of educational and communication needs (Dowdall et al., 2002) of the Greek Deaf Community.

Closing, we may notice that a children's dictionary (Kourbetis & Efthimiou, 2003) has already been developed, following the release of the NOEMA dictionary, which will provide further linguistic material for educational applications addressing early primary school needs.

Acknowledgments

The authors wish to acknowledge the assistance of all groups involved in the implementation of the proposed methodologies: terminologists, computational linguists, computer scientists, as well as the panel of GSL consultants and Deaf individuals, acting as GSL specialists and computer skills teachers.

More specifically, the authors wish to acknowledge the assistance of Dr. V. Kourbetis, S. Antonakaki, D. Kouremenos and V. Sollis.

This work was partially supported by the national project grants NOEMA (98 AMEA 11) and PROKLISI-EQUAL.

References

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

Dowdall, J., Hess, M., Kahusk N., Kaljurand, K., Koit, M., Rinaldi F. & Vider, K. (2002). Technical Terminology as a Critical Resource. In Proceedings of the Third International Conference on Language

Resources and Evaluation (LREC 2002), (pp. 1897--1903), Las Palmas de Gran Canaria, ELRA.
Efthimiou, E., Sapountzaki, G., Carpouzis, C., Fotinea, S-E. (2004). Developing an e-Learning platform for the Greek Sign Language. Lecture Notes in Computer Science (LNCS), Springer-Verlag Berlin Heidelberg (in print).
Efthimiou, E. & Katsoyannou, M. (2002). NOEMA: a Greek Sign Language Modern Greek bidirectional dictionary. Modern Education Vol. 126/127, 2002, 115--118 (in Greek).
Efthimiou, E. & Katsoyannou, M. (2001). Research issues on GSL: a study of vocabulary and lexicon creation. Studies in Greek Linguistics, Vol. 2 Computational Linguistics, 42--50 (in Greek).
Gee, J. & Goodhart, W. (1985). Nativization, Linguistic Theory, and Deaf Language Acquisition. Sign Language Studies, 49, 291--342.
Gougenheim, G. (1958). Dictionnaire fondamental de la langue française. Didier.
Hellenic National Corpus, Statistics. ILSP (Institute for Language and Speech Processing) (1999). Available at: http://www.xanthi.ilsp.gr/corpus/
Johnston T. & Schembri, A. (1999). On defining Lexeme in a Signed Language. Sign Language and Linguistics 2:2, 115--185.
Kourbetis, V. & Efthimiou, E. (2003). Multimedia Dictionaries of GSL as language and educational tools. Second Hellenic Conference on Education, Syros (in Greek), (in print).
Logiadis, N. & Logiadis, M. (1985). Dictionary of Sign Language. Potamitis Press (in Greek).
Longman Dictionary of Contemporary English.
Mueller, J., Bock, H. & Georgiakaki, M. (1995). Grundwortschatz Deutsch. Langenscheidt.
NOEMA Project (2001). Deliverable titled “Final implementation report”. ILSP, 6/2001 (in Greek).
Otman, G. (1996). Les représentations sémantiques en terminologie, Paris, Masson.
Prillwitz et al. (1989). HamNoSys. Version 2.0. Hamburg Notation System for Sign Language. An Introductory Guide. Broschur / Paperback (ISBN 3-927731-01-3).
PROKLISI Project: WP9: Development of a GSL based educational tool for the education of people with hearing impairments, Deliverable II: GSL linguistic material: data collection methodology. ILSP, May 2003.
Rey, A. (1995). Essays on terminology, Amsterdam/ Philadelphia, John Benjamins.
Rey, A. (1996). Beyond terminology.In Somers, H. (ed.), Terminology, LSP and Translation. Studies. in Language Engineering in Honour of Juan C. Sager (99--106), Amsterdam, John Benjamins.
Rousseau, L.-J. (1983). La définition terminologique. In Duquet-Picard, D. (ed), Problèmes de la définition et de la synonymie en terminologie: Actes du Colloque international de terminologie (35--46), Québec, GIRSTERM.
Sager, J. C. (1990). A Practical Course in Terminology Processing, Amsterdam, John Benjamins.
Sager, J. C. (1994). Terminology: Custodian of knowledge and means of knowledge transfer. Terminology 1/1, 7--16.
Sowa, J. F. (2000). Knowledge Representation: Logical, Philosophical, and Computational Foundations, Pacific Grove, Brooks Cole Publishing Co.
Tennant R.,A. & Gluszak Brown, M. (1998). The American Sign Language Handshape Dictionary. Clerc Books, an imprint of Gallaudet University Press.
Wright, S.-E. & Strehlow, R. A. (1995). Standardizing and harmonizing terminology: theory and practice, Philadelphia, American Society for Testing Materials.
Wilcox, S., Scheibmann, J., Wood, D., Cokely, D. & Stokoe, W. (1994). Multimedia dictionary of American Sign Language. In Proceedings of ASSETS Conference, Association for Computing Machinery, (pp. 9--16).

Λήμμα	Αντώνυμα	Συνώνυμα	Προέλευση	Γραμ	Γνωστικός Τομέας Α	Γνωστικός Τομέ	Παιδικό	Δόθη
ναυάγιο			KOI +		ΣΥΓΚΟΙΝΩΝΙΑ - ΜΕΤΑΦ		☐	
νότιος, νότια	βόρειος, βόρι		KOI +		ΠΕΡΙΒΑΛΛΟΝ		☐	
ρίζα			KOI +		ΦΥΤΑ		☐	
μπουκάλι			KOI +		ΤΕΧΝΟΛΟΓΙΑ - ΕΡΓΑΛΕ		☑	
μήκος, μάκρος		μάκρος (!!!)	KOI +		ΧΡΟΝΟΣ - ΧΩΡΟΣ - ΔΙΑΣ		☐	
μονός			KOI +		ΕΝΟΤΗΤΑ - ΔΙΑΙΡΕΣΗ -		☐	
αποδεικνύω, απόδ			KOI +		ΥΠΑΡΞΗ - ΤΑΥΤΟΤΗΤΑ		☐	
δηλητηριάζω, δηλ			KOI +		ΣΩΜΑ - ΕΜΦΑΝΙΣΗ - ΑΙΣ		☐	
ξεχωρίζω			KOI +		ΥΠΑΡΞΗ - ΤΑΥΤΟΤΗΤΑ		☐	
δεξιός, δεξιά	αριστερός, αρ		KOI +		ΧΡΟΝΟΣ - ΧΩΡΟΣ - ΔΙΑΣ		☐	
μόνο, μονάχα		μονάχα (!!!)	KOI +		ΕΝΟΤΗΤΑ - ΔΙΑΙΡΕΣΗ -	ΓΡΑΜΜΑΤΙΚΗ	☐	
νησί			KOI +		ΠΕΡΙΒΑΛΛΟΝ		☐	
ωραίος	άσχημος	όμορφος	KOI / παρ+ / Δη		ΣΩΜΑ - ΕΜΦΑΝΙΣΗ - ΑΙΣ		☐	
καινούριος	παλιός		KOI / παρ+		ΧΡΟΝΟΣ - ΧΩΡΟΣ - ΔΙΑΣ		☐	
ακόμα, ακόμη			KOI / παρ+		ΓΡΑΜΜΑΤΙΚΗ		☐	
δεν, μην			KOI / παρ+		ΓΡΑΜΜΑΤΙΚΗ		☐	
ακούω			KOI / παρ+		ΣΩΜΑ - ΕΜΦΑΝΙΣΗ - ΑΙΣ		☑	
μεγαλώνω	!!!		KOI / παρ+		ΧΡΟΝΟΣ - ΧΩΡΟΣ - ΔΙΑΣ	ΧΡΟΝΟΣ - ΧΩΡΟΣ	☐	
καθόλου, διόλου		διόλου (!!!)	KOI / παρ+		ΕΝΟΤΗΤΑ - ΔΙΑΙΡΕΣΗ -		☐	
κάθομαι			KOI / παρ+		ΕΝΕΡΓΕΙΑ - ΑΔΡΑΝΕΙΑ -		☑	
με	χωρίς		KOI / παρ+		ΓΡΑΜΜΑΤΙΚΗ		☐	
δένδρο			KOI / παρ+		ΦΥΤΑ		☑	
άκρη, ακμή	μέση		KOI / παρ+		ΧΡΟΝΟΣ - ΧΩΡΟΣ - ΔΙΑΣ		☐	
και			KOI / παρ+		ΓΡΑΜΜΑΤΙΚΗ		☐	

Record: 1800 of 2823

Figure 1: Part of the GSL basic vocabulary DB; the 3rd column from left provides information as regards original (co-) appearance of lemmas in source lists.

Λέξη αναζήτησης	Video	Ερμηνεία	Αντώνυμα	Συνώνυμα	Γνωστικός Τομέας
επάνω	1981pano-sto.avi	πάνω σε			ΤΟΠΙΚΕΣ ΕΝΝΟΙΕΣ
επάνω	1980epano.avi	πάνω από			ΤΟΠΙΚΕΣ ΕΝΝΟΙΕΣ
επαρχία	1099eparxia.avi	επαρχία	1513protevusa.avi		ΚΡΑΤΟΣ - ΔΙΟΙΚΗΣΗ
επείγον	905epigon.avi	επείγον			
επειδή	319dioti.avi	διότι, επειδή			
επεισόδιο	320episodio.avi	επεισόδιο, τσακωμός, φασαρία			
έπειτα	1982epita.avi	έπειτα, μετά	2132prin.avi, 2135proigumenos	2064meta.avi	ΧΡΟΝΙΚΕΣ ΕΝΝΟΙΕΣ
επεξεργάζομαι	321epexergazome.avi	επεξεργάζομαι			
επί	1647epi.avi	επί (σύμβολο πολλαπλασιασμού)			
επί τοις εκατό	1100epitisekato.avi	επί τοις εκατό			ΟΙΚΟΝΟΜΙΑ
επιβάλλω	2338epivalo.avi	επιβάλλω			
επιβάτης	3322epivatis.avi	επιβάτης (άντρας ή γυναίκα)		3323epivatis2.avi	ΣΥΓΚΟΙΝΩΝΙΕΣ - ΜΕΤΑΦΟΡΕΣ
επιβάτης	3323epivatis2.avi	επιβάτης (άντρας ή γυναίκα)		3322epivatis.avi	ΣΥΓΚΟΙΝΩΝΙΕΣ - ΜΕΤΑΦΟΡΕΣ
επιγραφή	776epigrafi.avi	επιγραφή			
επιδιορθώνω	531sintiro.avi	επιδιορθώνω, συντηρώ			ΡΗΜΑΤΑ ΕΝΕΡΓΕΙΑΣ
επιδοκιμάζω	2448penevo.avi	επιδοκιμάζω		2451epeno.avi	
επίδραση	323epiroi.avi	επιρροή, επίδραση			
επίθεση	1101epithesi.avi	επίθεση	996amina.avi		

Record: 977 of 3197

Figure 2: Part of the GSL basic vocabulary DB; synonym and antonym association to video-lemmas.

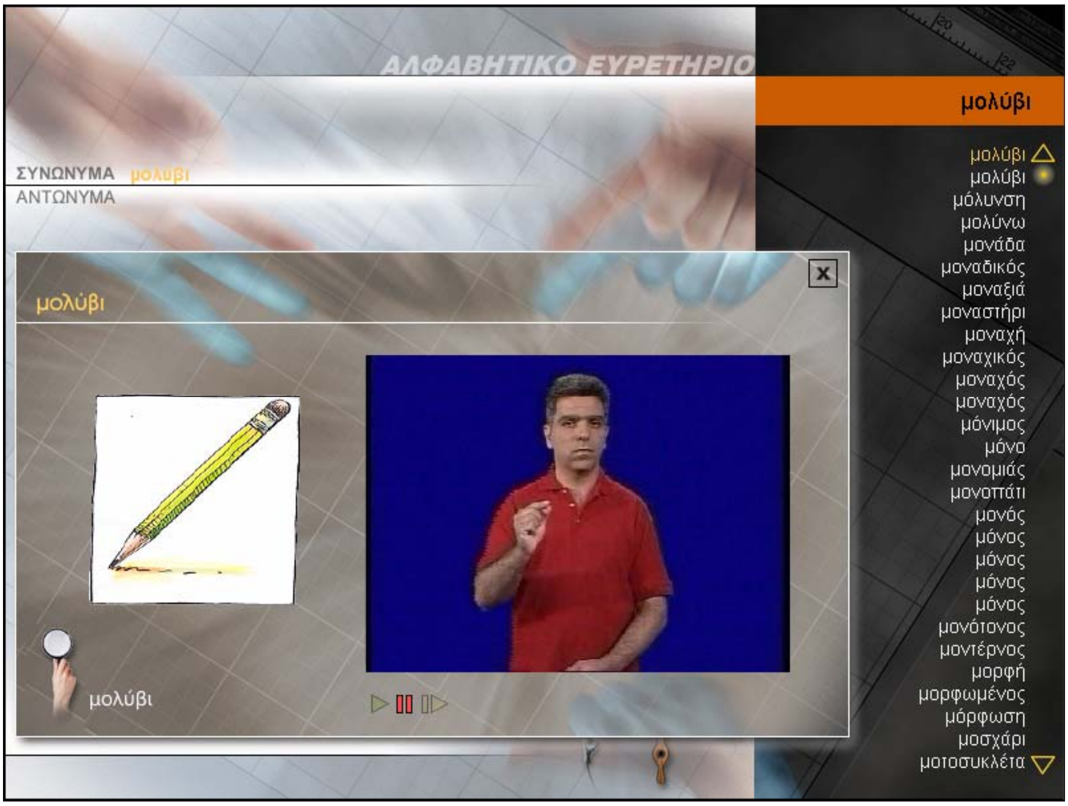


Figure 3: Synonym/antonym screen incorporated in alphabetical search capability.

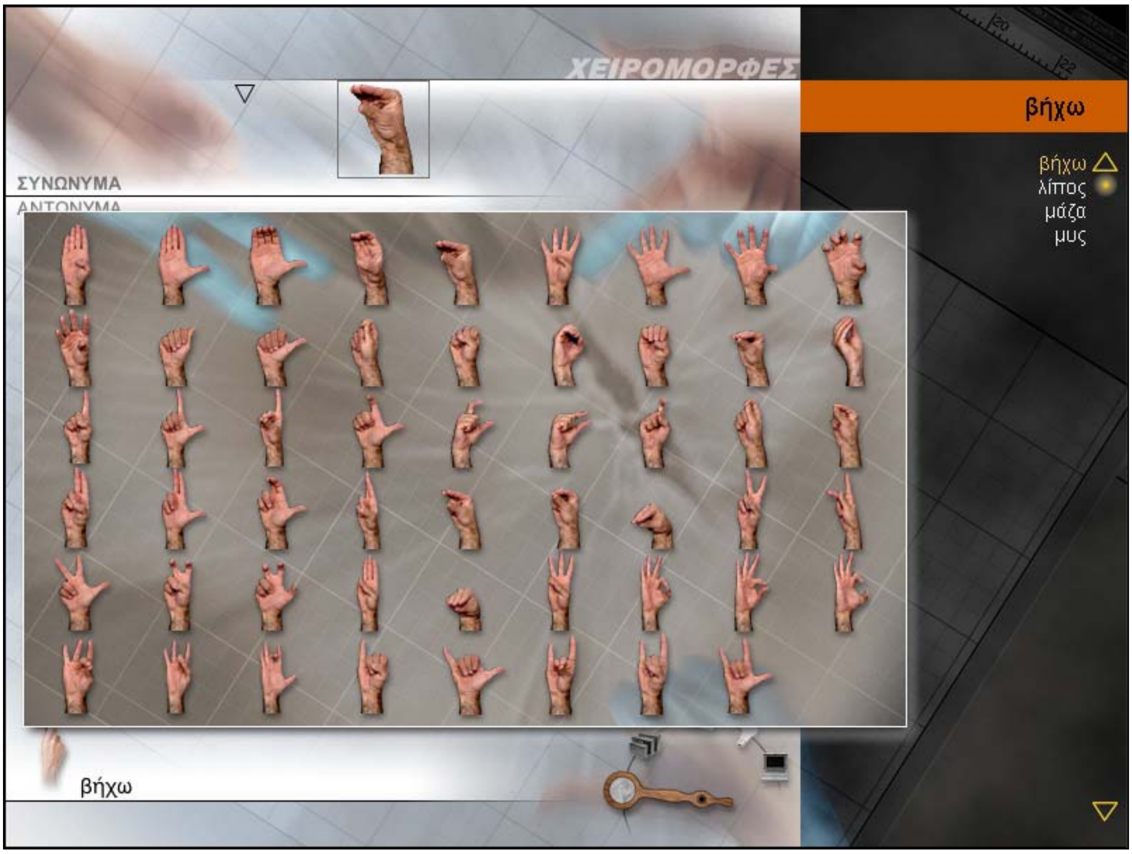


Figure 4: Lemma search by handshape in the GSL – Modern Greek basic vocabulary dictionary.

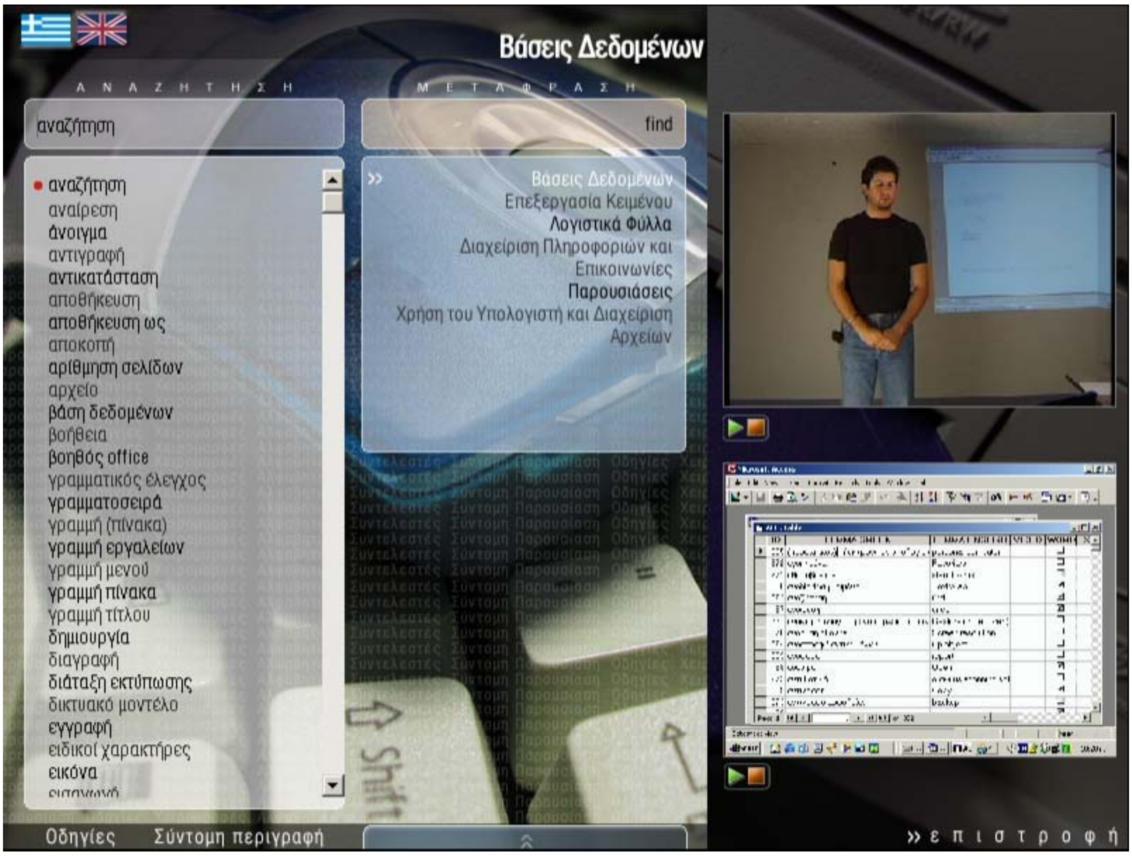


Figure 5: Computer-skills term dictionary: alphabetical search screen.

ID	LEMMA GREEK	LEMMA ENGLISH	VIDEO	WORD	EXCEL	ACCESS	INTERNI	POWERI
572	αγαπημένα, σελιδοδείκτης	favorites, bookbark		☐	☐	☐	☑	☐
271	αθόρυβος ιός	stealth virus		☐	☐	☐	☐	☐
56	αναδίπλωση κειμένου	text wrap		☑	☑	☐	☐	☑
363	αναζήτηση	find		☑	☑	☑	☑	☑
57	αναίρεση	undo		☑	☑	☑	☐	☑
661	ανάκτηση (διαγραμμένου αρχείου ή φακ	restore, recover (dele		☐	☐	☐	☐	☐
670	ανάλυση οθόνης	screen resolution		☐	☐	☐	☐	☐
364	αναστροφή αντικειμένων	flip object		☐	☐	☐	☐	☑
259	αναφορά, έκθεση	report		☐	☑	☑	☐	☐
58	άνοιγμα	open		☑	☑	☑	☐	☑
272	αντιβιοτικό	antivirus scanning sof		☐	☐	☐	☐	☐
59	αντιγραφή	copy		☑	☑	☑	☑	☑
273	αντίγραφο ασφαλείας	backup		☐	☐	☐	☐	☐
62	αντικατάσταση	replace		☑	☑	☑	☐	☑
371	αντίτυπα	copies		☑	☑	☑	☑	☑
274	αποδιαμόρφωση	demodulation		☐	☐	☐	☐	☐
63	αποθήκευση	save		☑	☑	☑	☐	☑
64	αποθήκευση ως	save as		☑	☑	☑	☑	☑
686	αποθηκευτικός δίσκος	storage disk		☐	☐	☐	☐	☐
375	αποκάλυψη	show		☐	☐	☐	☑	☑
376	αποκοπή	cut		☑	☑	☑	☐	☑
377	απόκρυψη	hide		☐	☐	☑	☐	☑
268	αποσύνδεση	log off		☐	☐	☐	☐	☐

Record: 1 of 388

Figure 6: Association of lemma to sub-area of field in computer-skills terminology DB.

From Computer Assisted Language Learning (CALL) to Sign Language Processing: the design of e-LIS, an Electronic Bilingual Dictionary of Italian Sign Language and Italian

Chiara Vettori, Oliver Streiter, Judith Knapp

Language and Law
EURAC; European Academy of Bolzano
Viale Druso/Drususallee 1, 39100 Bolzano/Bozen, Italy
{cvettori;ostreiter;jknapp}@eurac.edu

Abstract

This paper presents the design of e-LIS (Electronic Bilingual Dictionary of Italian Sign Language (LIS) and Italian), an ongoing research project at the European Academy of Bolzano. We will argue that an electronic sign language dictionary has to fulfil the function of a reference dictionary as well as the function of a learner's dictionary. We therefore provide an analysis of CALL approaches and technologies, taking as example the CALL systems ELDIT and GYMN@ZILLA developed at the European Academy of Bolzano too. We will show in how far these approaches or techniques can be ported to create an electronic dictionary of sign languages, for which system components new solutions have to be found and whether specific modules for the processing of sign languages have to be integrated.

1. Introduction: Dictionaries of LIS

Around 50.000 people in Italy are deaf. The first language of the majority of them is LIS, Lingua Italiana dei Segni (Italian Sign Language), but there is also an undetermined percentage of oralist deaf people. LIS is also acquired as a second or third language by hearing family members, teachers, interpreters and logopedics, amounting to about 170.000 people using LIS, in various degrees of language competence. Unfortunately, the quality and accessibility of LIS-courses and supporting material (dictionaries, text books, and videos) lack behind the actual need. Moreover, the official support does not meet the high standards of other countries and does not comply with international recommendations, e.g. Recommendation 1598 (Council of Europe 2003), which advice, among others, to broadcast television programs in sign language, to utilize new technologies for teaching sign languages and to include sign languages as a valid academic qualification. It is most likely that such status quo also depends on the position of the Italian government which has not yet officially recognized LIS.

As for LIS dictionaries, the vast majority of them are paper based ones, e.g. Radutzky 1992 (752 signs, 2500 sign meanings); Angelini et al. 1991 (400 signs). The paper format, however, cannot obviously account for the possibility of describing the three-dimensional complexity of each sign. A first, significant attempt in Italy to exploit new technologies to approach sign languages in an innovative and more proficient way, was made by the team of Cooperativa Alba. Its members have opened an Internet portal for LIS (DIZLIS) that now features more than 1000 video-filmed signs, which represent a respectable size for a sign language dictionary, cfr. Sternberg 1987 (3300 signs), Stewart et al. (2500 signs). Italian serves as vehicular language and dictionary index.[1] The advantage of this presentation of signs over the schematic and static drawing in paper dictionaries is evident and has motivated similar projects in other countries.

2. Towards e-LIS

Most sign language dictionaries form a hybrid between a reference dictionary and a learner's dictionary. This often occurs because sign language is implicitly considered as the second language of a "learner's dictionary" de-facto created for the needs of hearing people. At the same time these lexicographic works pretend to fulfil the function of a reference dictionary of the involved sign language, only in virtue of the presence of drawings and photos representing different signs. "*A major feature of such dictionaries is the absence of definitions, it being assumed that each sign would have exactly the same meaning(s) as the written word with which it is linked*" (Brien 1997). This sort of production treats signs as equivalents of the words of a spoken language and neglects the complexity, the dignity of sign language and its peculiarities in semantics and syntax. Lexical units in a sign language differ in a number of important features from the translational equivalents in the spoken language. These are:

- the referential extension, i.e. which objects, states and events are referred to by a word,
- the conceptualization, e.g. as *Abendstern*, *Morgenstern* or *Venus* (Frege 1892),
- the micro-syntax, e.g. the derivational history of a word from its bases via compounding or derivation to its final form,
- the stability with which they belong to a word class (nouns vs. verbs),
- the lexical relations they maintain, e.g. expressed as lexical functions (Melc'uk 1974) and
- the affiliation of a word to a word class which does not exist in the other language, e.g. classifiers in sign language or functional prepositions in the spoken language.

[1] http://www.dizlis.it

As LIS is an autonomous language and not a mere visual representation of Italian, we designed a dictionary which describes two systems at the same time, the Italian and the LIS one, and which can also build a bridge between them through a sort of "translating interface". In this perspective, accepting Stokoe's description of what he calls "serious dictionaries" (Brien 1998), we are greatly motivated to focus on the definition of sign meanings that could reveal much of the deaf culture.

This accommodates for two distinct user groups. (a) Hearing Italian people who study LIS and who will start with an Italian query term in an Italian environment (Italian definitions, explanations etc.); (b) LIS-signers looking for a sign and who should have the possibility to formulate query terms in LIS and have a LIS environment.

In order to assure the description of sign language in the sign language itself, therefore accounting for the specificity of this linguistic code[2], appropriate modes of rendering it into a Web-interface are required. One unexplored way of providing signs' definitions could be realized through the adoption of SignWriting (Rocha Costa & Pereira Dimuro 2003). In contrast to filmed definitions, in fact, SignWriting renders the definitions, explanations and menu buttons searchable (Aerts et al. 2004, Rocha Costa et al. 2004) and hyperlinkable. Words contained in a definition may thus be linked to lexical entries, which feature, as main component, the filmed sign.

3. ELDIT

One of the tools we already count on and from which we intend to develop the e-LIS dictionary is ELDIT, an electronic learners' dictionary for German and Italian. Inspired by the lexicographic research started in the '50s and according to recent psycholinguistic and didactic theories (Aitchison 94, Kielhöfer 96), it covers a limited vocabulary consisting of approximately 3.000 words for each language. It also stores a large set of information for each word entry and highly interlinked information pieces.

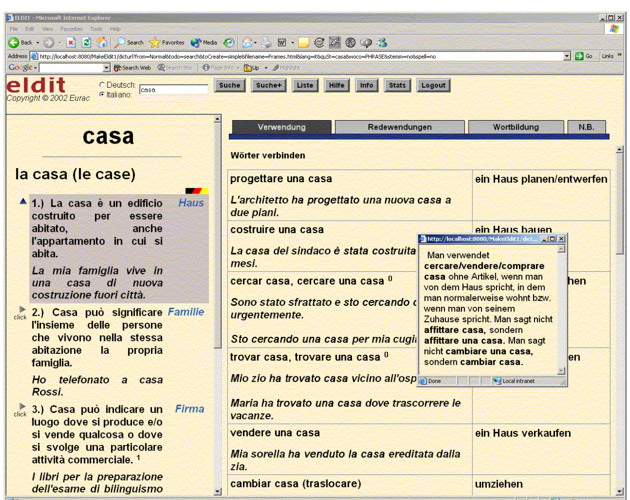


Figure 1: Dictionary entry for the Italian word "casa" (house) in ELDIT.

Figure 1 shows a screenshot of the dictionary entry for the Italian word "casa" (Engl. "house"). The left-hand frame shows the different word meanings. Each meaning is described by a definition, a translation, and an example. The right-hand frame shows additional information, which depends on the selected word meaning. The collocation tab lists the most frequent collocations along with their translation and an illustrative example. In the semantic field tab word relations (such as synonymy, antonymy, etc.) are illustrated in graphs for the learner. Verb valency is explained using colours and movable elements. Adopting a comparative approach, ELDIT also stresses specific differences between the Italian and the German language. Such differences are indicated by footnote numbers. Last but not least, each word used in the system (e.g. in the definitions or in the example sentences) is annotated with lemma and part-of-speech and is linked to the corresponding dictionary entry, which facilitates a quick dictionary access for unknown words.

4. GYMN@ZILLA

A further interesting way of facing LIS and Italian is represented by Gymn@zilla, a browser-like application which integrates existing educational and non-educational modules in a new didactic environment. Gymn@zilla allows to access documents from the Internet and to convert its text into an easy reader text, a glossary and a completion exercise.

Gymn@zilla is used like any browser. The program accesses a web-page, identifies its language and encoding and performs a simple word-form stemming of text. The stemmed words and expressions are then linked to their respective translations in external dictionaries. The linked lemma is marked up as html tool-tip to provide an immediate translation aid even without following the external link.

Clicking on a word triggers two actions. First, the complete explanations of the external lexicon are opened. Second, the word, its context and its translation are added to a personal glossary. The learner can edit the vocabulary in his personal dictionary and use it for intentional vocabulary acquisition, as opposed to incidental vocabulary acquisition by annotated reading of the web-page. Last, the learner can create interactive quizzes from the personal glossary, for which Gymn@zilla automatically offers inflected, uninflected and misspelled forms to fill the gaps. Gymn@zilla handles a number of language pairs, some going from a spoken language to a Sign Languages (e.g. English=>ASL, c.f. Figure 2). Through a triangulation of the translation dictionaries (e.g. Italan => English => ASL) we will give Gymn@zilla new dimensions of usage.

[2] Cfr: Les Signes de Mano
http://www.ivtcscs.org/media/mano.htm

Figure 2: Annotated reading with Gymn@zilla.

5. e-LIS Architecture

Hence it becomes obvious, even after this schematic analysis, that an electronic dictionary of sign language can be much more than a list of search indices, each hyperlinked to a video file. The search will start with an Italian key word or a LIS key word entered in SignWriting yielding a parallel list of all matching lemmas and collocations in Italian-LIS (SignWriting), similar to *LEO*[3], developed by the Munich University of technology, and *bistro*[4], developed at the European Academy Bolzano. Clicking on a word or an expression makes this a search term, possibly inverting the direction of search. As in *bistro*, additional links will lead to the monolingual lexical entries.

The Italian entry will be close to its current form in ELDIT, which might be profitably reused for developing e-LIS (c.f. Figure 1). Link texts to related entries in LIS will be rendered in SignWriting. The LIS entry will feature the filmed representation of the LIS sign. All definitions and explanation in the LIS entry will be in LIS, rendered in SignWriting. As in the Italian entry, each sign will be hyper-linked to the corresponding LIS entry. Lexical functions, e.g. classifiers, collective nouns (antelope => herd, ant => army) etc. will be realized as hyperlinks to entries as well, as well as the backward relation. Example sentences, collocations, idioms in LIS which do not have a proper lexical entry will be directly linked to the filmed sign presentation. As for the video approach, we will draw on the materials already developed for the site DIZLIS by the Cooperativa Alba.

[3] http://dict.leo.org/

[4] http://www.eurac.edu/bistro

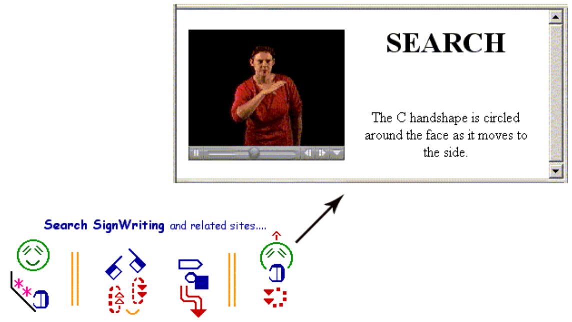


Figure 3: SignWriting in combination with Sign Language, a vision for the e-LIS system.

Beside this kind of inner metalinguistic description, we won't forget the peculiar needs of Italian speaking learners of LIS who will presumably not be able to read SignWriting and prefer videos of signs. For these users, as well as for signers studying Italian, Gymn@zilla can be easily invoked with its habitual functions:

- Italian words will be rendered as easy reader through video films or SignWriting
- SignWriting will be rendered as easy reader through video films or Italian
- personal word lists can be constructed
- completion test can be started at any time (in Italian and SignWriting)
- Texts Italian and SignWriting located in the WWW can be smoothly integrated into e-LIS, with proposed or freely selected texts, in order to allow the first steps outside the e-LIS environment. In case of any doubt, Gymn@zilla will take the user always back to e-LIS to provide the necessary explanations.

In addition, the analysis of possible difficulties LIS-signers encounter in studying Italian (Taeschner et al., 1988; Fabbretti 2000; etc.) suggests another usage a sign-language dictionary could be put to. We intend to supply the dictionary with an apparatus of contrastive grammatical macros in analogy to the footnote numbers in ELDIT. These macros are triggered whenever a lexical entry contains critical features, e.g. semantically weak prepositions such as "di" ("of") which cause translation difficulties for signers while writing in Italian, differences in word order etc. The lexical material of the entry and its parallel counterpart (in LIS or Italian) will be inserted into the macro and rendered from the point of view of the actual entry, yielding a comparative and synoptic descriptions of challenging grammatical aspects of the two languages compared with the lexemes of the current entry. Also in this perspective, the use of SignWriting could be particularly useful because it permits to parcel two equivalent strings in sign language and Italian and to interrelate the single syntagms/parts thus immediately showing the similarities and differences of the two systems with the aid of colours (for the corresponding elements) and explanations in sign language in case of differences.

6. Conclusions

We have presented so far the rationale of e-LIS, Electronic Bilingual Dictionary of Italian Sign Language (LIS) and Italian. A short analysis of existing Sign Language projects and of several CALL projects that have been carried out in the past years at the European Academy Bolzano has revealed that an electronic dictionary of sign language can be much more than a simple list of search indices, each hyperlinked to a video file.

While reusing some tools and options of the Italian-German dictionary ELDIT and enriching them through the many didactic functions provided by Gymn@zilla, a browser that converts Internet texts into easy reader ones, we will develop a new type of sign language dictionary.

We hope that our system might contribute to a research area that up to now has been quite neglected in Italy and that it could contribute to and accelerate the process which will lead Italian government to the official acknowledgement of LIS.

7. References

Aerts, S., Braem, B., van Mulders, K. De Weerdt, K., 2004. Searching SignWriting Signs*, this volume*.

Angelini, N., Borgioli, R., Folchi, A. and Mastromatteo, M., 1991. *I primi 400 segni. Piccolo dizionario della Lingua Italiana dei Segni per comunicare con i sordi.* Firenze: La Nuova Italia.

Brien, D. and Collins, J., 1997. The CD-ROM Dictionary of Deaf Community and Culture: Some Issues in the Creation of Sign Language Dictionaries. Paper presented at Empower'97: International Conference on Deaf Education. http://www.ssc.mhie.ac.uk/archive/proceedings/briencollins.html

Fabbretti, D., 2000. L'italiano scritto dai sordi: un'indagine sulle abilità di scrittura di sordi adulti segnanti nativi. *Rassegna di psicologia*, 1 (XVII):73-93.

Frege, G. 1892. *Über Sinn und Bedeutun,* Ztschr. f. Philosophie und philosophische Kritik, NF 100:25-50.

Melc'uk, I., 1974. *Opyt teorii lingvističeskix modelej 'Smysl-Tekst'*. Moscow, Nauka.

Radutzky, E. (ed), 1992. *Dizionario bilingue elementare della Lingua Italiana dei Segni*. Rome: Edizioni Kappa.

Rocha Costa, C. A. and Pereira Dimuro, G., 2003. SignWriting and SWML: Paving the Way to Sign Language Processing. Workshop *Traitement automatique des langues minoritaires et des petites languges, TALN 10e conference, tome2,* 193-202.

Rocha Costa, C. A. and Pereira Dimuro, G., Balestra de Freitas, J., 2004. A sign matching technique to support searches in sign language texts. *, this volume*.

Sternberg, M.L.A. (ed), 1987. *American Sign Language Dictionary*. New York: Harper & Row, New York.

Taeschner, T., Devescovi, A., Volterra, V., 1988. Affixes and function words in the written language of deaf children. *Applied Psycholinguistics*, 9:385-401.

19 th Century Signs in the Online Spanish Sign Language Library: the Historical Dictionary Project

Rubén Nogueira & Jose M. Martínez
University of Alicante
Ap.Correos 99.-E-03080 Alicante (Spain)

ruben.nogueira@ua.es, jmanuel.martinez@ua.es

Abstract

This paper will illustrate the work made in the Sign Language Virtual Library (http://www.cervantesvirtual.com/portal/signos), a project aimed at people interested in Spanish Sign Language; and specially, at its main users, Deaf people. It is organised into six different sections: Literature, Linguistics, researchers forum, Deaf culture and community, bilingual-bicultural education and didactic materials. Each section contains different publications related to the above mentioned areas. Moreover, in this web you will also find an innovation, since every publication includes a summary in Spanish sign language. Two sections will be described: The Historical Dictionary published by Francisco Fernandez Villabrille and the Alphabetical Writing Lessons. Our intention is showing a full functional version of the applications described on the paper.

1. Introduction

All languages are natural phenomena for the people who constantly use them to communicate. This is also the case of deaf people and Sign Languages. For different reasons, these languages have lacked the consideration enjoyed by oral languages. In fact, certain misconceptions – that sign languages were artificially created, that they are universal expressions, that they are merely mimetic or that they were created to replace oral languages – still exist, although not as much in the field of linguistics as in the rest of society.

Several psycholinguistic studies (Bellugi, Klima & Siple, 1975) have indicated the natural acquisition of these languages, as well as the various stages or phases of development that deaf children must go through when learning to sign; these stages are similar to those established for hearing children who learn an oral language and may even occur earlier during the process. (Folven & Bonvillian, 1991; Juncos et al., 1997).

Using gestures to communicate is inherent to human beings; in this case, we could say that it is a universal tendency. However, the codes applied by users of sign languages relate to different cultural and linguistic patterns, in their phonetics, morphology, syntax… It must be emphasized as far as possible that Sign is a language like any other, because – although it may seem otherwise – most hearing people have never had contact with deaf people and are totally unaware of their reality.

We will dwell here on a very important issue concerning the study of any language: how has it been modified through time?

Observing these changes, not only throughout time, but in accordance with the universal tendency of evolution of languages, it is now possible to speak of the linguistic evolution of Spanish Sign Language as the normal historical evolution of a language.

Publications available in Spain (Hervás, 1795; Bonet, 1620; Ballesteros, 1845), among others, show that, thanks to the efforts of Brother Ponce de León, the education of Deaf people started a long time ago in Spain. At the time, Deaf people were taught to acquire an oral language: Spanish or Latin, and education was a privilege available only to a few, mainly children of the nobility, who, in exchange, favoured the clergy financially. However, the following facts must be taken into account:

- until then it was commonly believed that Deaf people had no understanding or language, this way of thinking changed thanks mainly to Ponce de León.
- Ponce de León invented the manual alphabet and this is one of the first proofs of the visual-gestural characteristics of the language of Deaf people and their means of communication.

Thanks to these publications, now included in our website: Biblioteca de Signos, The Online Spanish Sign Language Library, we know of the existence of a tradition of deaf education in Spain, begun by Pedro Ponce de León and Juan Pablo Martínez Carrión. This tendency was known in Europe as 'The Spanish School' or 'The School of the Art of Teaching The Mute To Talk'.

The word 'mute' appears in almost every one of these books; the first to mention the difference between 'deaf-mute' and 'deaf' is Lorenzo Hervás, in 1795. The Spanish School was very important until the end of the 18th century.

Two of these very important books deserve particular mention: *La Escuela Española de Sordomudos o Arte para enseñarles a escribir y hablar el idioma español* (The Spanish School of Deaf-Mutes or The Art of Teaching Them To Write and Speak Spanish) written by the linguist Lorenzo Hervás, an important expert in linguistic typology, and the *Diccionario usual de mímica y dactilología* (Dictionary of Usage of Miming and Dactylology), by Professor Francisco Fernández Villabrille, which is the starting point for the *Diccionario Histórico de la LSE* project (The Historical Dictionary of Spanish Sign Language).

2. The Historical Dictionary Project

More than a century has passed since the first Spanish Sign Language dictionary was published by Francisco Fernández Villabrille in 1851. His work gives us a date for the formation and consolidation of this language.

The project introduced here began when we were given the opportunity of offering this text, with its translation into Sign Language, on the Internet, through the *Biblioteca Virtual Miguel de Cervantes* (Miguel de Cervantes Virtual Library), an ambitious project for the digital publication of Hispanic culture, with over twelve thousand digitalised books, in their respective sections. In one of these sections, The Sign Language Library, we translated all these texts into Sign Language, using video combined with other elements, as we will explain in detail below. Our intention goes beyond leaving a testimony of the signs used by the Deaf in the 19^{th} century and, accordingly, when we finished presenting the book and the signs in LSE, we wanted to round off our work with later dictionaries and current signs.

In order to prepare the Historical Dictionary project, the team, composed of deaf people, Sign Language interpreters and linguists with expertise in Sign Language, first thoroughly revised the roughly 1500 signs contained in Villabrille's dictionary. In doing this, we studied the most descriptive phonologic components in the dictionary and also looked for similarities and differences with the current phonological system of SSL. There are cases of disappearance, modification or addition of phonemes and other interesting phenomena such as variation of locations, hand shapes, movements and fluency of components (assimilation) (Woodward, 1975; Frishberg, 1975). Morphologically, we could mention resources used for gender inflexion, number, person, tense, etc., through repetition or a certain use of direction, among others. Syntactically speaking, we have less information but we can still discuss structure and the rules for combining signs, which have, on occasions, undergone changes and on others, stayed the same (incorporating syntactic interferences from related languages, such as oral Spanish). Semantically, we can see a transformation in the natural evolution of the linguistic situations in which the language is used, which implies that Sign Language has the capacity to develop and evolve over time, broadening/restricting meanings, borrowing from related languages, etc.

During this revision we pondered, for instance, whether the signs given in the document are still in use or whether they have disappeared; we also analysed whether or not the signs are mimetic representations. Some signs that used to be mimetic ('agile' - AGIL, 'to fall' - CAER...) no longer are; others were originally arbitrary ('cat' - GATO, 'to make' - HACER). No language is systematically mimetic. For example, let's take old sign for 'to fall' – CAER, described in the dictionary as an imitation with the body representing the movement of the action. This representation cannot be considered linguistic. Deaf people are conscious of whether they are miming or using a linguistic sign, but this matter is somewhat more complex for researchers, as the difference is not always sufficiently clear.

We have tried to distinguish conventional from non-conventional, although we must take into account the fact that we are dealing with a language with no register or standardization rules.

Once we had finished this analysis and agreed on the signed production of each sign (taking into account that our information was occasionally insufficient and, accordingly, could not be recorded), the next step was to record the almost 1400 signs in a professional video workshop, to later capture them, treat them digitally and upload them onto the web page, where users may watch the video simultaneously with a description of the sign given in Villabrille's work and the new Alphabetic Writing System for SSL, created by the University of Alicante.

We will now explain the technical matters relating to this project.

3. Technical Description of the Historical Dictionary

Some time ago, the people working on the Online Spanish Sign Language Library project expressed a wish to create an online version of the Historical Dictionary published by Francisco Fernandez Villabrille in 1851.

Applying some of the concepts previously used in the Online Spanish Sign Language Library, we arrived at the following design:

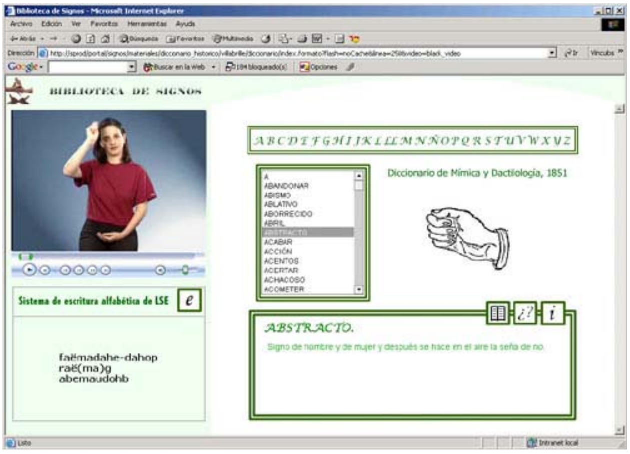


Figure 1: Historical Dictionary design.

First of all, the user chooses a letter of the alphabet, by simply clicking on it. Next, a list appears with all the words in the dictionary beginning with that letter, accompanied by an image with the corresponding historical sign. We now choose the word that we wish to consult from the list, and the definition of the word appears. Depending on the word selected, several informative icons may be activated. Additionally, the choice of a certain word, can initiate two more actions: a video reproduction of the signed representation of the selected word, and its written representation (SEA). This is, in broad strokes, the challenge issued to the computer science department of the Online Spanish Sign Language Library project.

Until then, we had a clear series of concepts regarding the integration of different technologies, such as the exchange of "messages" between a flash file embedded in a Web document, the document itself and a video file. However, although these concepts were clear, we still needed to solve the biggest problem: the local management of simple data bases; after all, we were talking about a dictionary. The solution was provided by Macromedia and the capacity of Flash to manage XML documents. We had now solved the database problem in local mode, thus avoiding unnecessary host requests. All the words, definitions and associated extra information would be generated using XML rules.

The application consists of 4 frames, one static and housing an image, the rest being dynamic, at least as regards their content. In other words, the contents of the frame would change without having to refresh. Let's take a look:

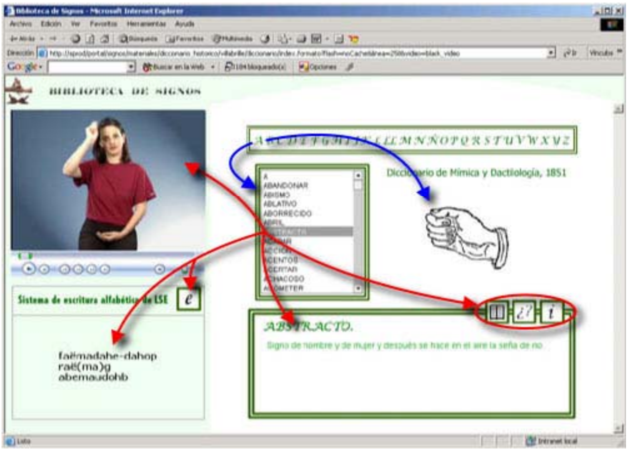


Figure 2: Arrows representing the exchange of messages

The core of the application would be located in the flash frame and would consist of a flash application embedded in the corresponding HTML document, plus all the JavaScript code included in the frame. This code manages all the messages sent by and to the flash file. We could say that it is the heart of the application, as it supports much of the graphic interface, manages and processes the data, and creates the messages arising from the preceding interpretation. These messages modify the contents of the video frame and the SEA frame. The changes that affect the flash application are, obviously, self-managed and require no messages.

Thanks to the design capacities of Macromedia Flash MX, it was relatively simple to transfer what we had devised on paper onto the screen (the interface).

The internal process could be summarized as follows:

1. A letter is selected.
2. The XML document associated with that letter is loaded in the application and an image is shown with the corresponding sign.
3. A word is selected.
4. The data accompanying the selected word is processed and, depending on the content, the following processes will be triggered:

a. The video with the signed representation of the word may start.

b. The written representation of the sign is shown in the Alphabetic Writing System (Sistema de Escritura Alfabética – SEA).

c. The definition of the word is shown.

d. The informative icons corresponding to the selected word are activated.

Processes a and b create a message that is interpreted by a JavaScript function (included in the HTML document where the flash application is embedded), which also modifies the video frame contents and the SEA frame contents. Processes c and d are internal to the flash application.

Looking at the current panorama, there are a number of web pages with online LS dictionaries out there: *Silent Thunder Animations , the ASL Browser, ASL University, A Basic Guide to ASL, Handspeak, Signhear ASL Dictionary.* However, in our opinion, these attempts do not give image the importance and the presence that it must have within the languages of signs. Our intention was to take full advantage of the capacity of image, using the latest video reproduction technologies used on the Internet (video streaming). By reinforcing the power of image using such a powerful tool as Flash, not only did we not turn our back on the possibility of using image on the Internet, but, rather, we increased its effects.

However, we wouldn't wish to give the impression that this is the entire scope of the project: we hope to make the dictionary a point of reference and therefore, we will continue to develop new versions of the application capable of supporting advanced term searching, a choice of different dictionaries, etc...

An essential part of our work is processing signed videos. The videos arrive at the laboratory in miniDV format, and they are captured and produced using Adobe Premiere (video editing software). After these stages, we now have a high quality version of the video, which must now be coded in order to adapt it to the size restrictions imposed by the Internet. The technology used at this stage is Windows Media Encoder, for the reason that commands can be included in the videos themselves (hypervideo). We are aware that we are working with proprietary software that may not be available to all users, but we considered that it was what best fit the needs of our portal. We do not reject the possibility of standardizing access to the portal, as far as possible, and it is our intention to approach these questions as soon as possible.

We will now complete the description of our work on the Online Spanish Sign Language Library through a practical example: the writing lessons.

4. Writing Lessons

In this case, the people working on the Online Spanish Sign Language Library project decided to include an online version of the writing lessons in the portal.

Applying concepts previously used in the Online Spanish Sign Language Library, we arrived at the following design:

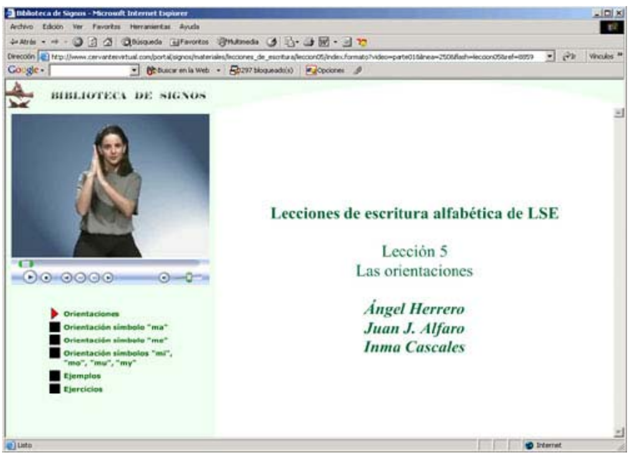


Figure 3: Writing Lessons design.

In this case, the user of the portal, having chosen a writing lesson, would be looking at a page similar in design to figure 3. The flash frame would show the title of the lesson and, immediately afterwards, the video would start up in the video frame. Based on the contents of the explanation, the flash application embedded in the flash frame will show different multimedia contents to reinforce the video explanation, thereby improving understanding of the lesson. It would also be necessary to include a menu indicating what part of the lesson the user is in and making it possible to select different sections from the selected lesson.

We simply had to apply the technology integration concepts mentioned in the Historical Dictionary project to give life to the proposed application.

As in the previous case, the application consists of 4 frames, one static and housing an image, the rest being dynamic, at least as regards their content. In other words, the contents of the frame would change without having to refresh. Let's take a look:

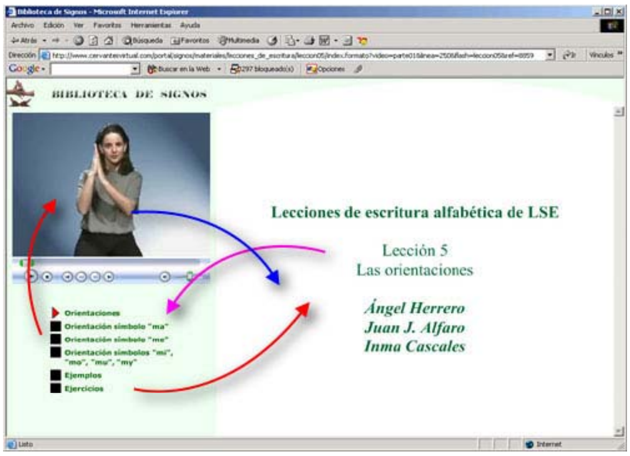


Figure 4: Arrows representing the exchange of messages

In this case, we used one of the special characteristics of clips encoded with Windows Media Encoder (the capacity to insert commands within the video itself). Once this is clear, it works as follows: we first take note of the exact moments at which the video clip must give way to a certain animations in flash frame. Next, the appropriate commands are inserted at those moments. Using the JavaScript functions included in the different frames, the desired effect was obtained. It is therefore obvious that the main body of the application is situated in the video (in Windows Media Video format) embedded in the video frame. However, it must not be forgotten that the dynamic frames can communicate with each other, and this is what provides the necessary flexibility to be able to design and adapt another flash application: the menu, from which the user can enter the lesson and select a section.

Looking at these cases, we can see the flexibility provided by the integration of several technologies in the process of developing multimedia applications. Each of the technologies can work alone, but they also provide the capacity to modify the state of the other technologies coexisting in the same surroundings. If we add the fact that these technologies are all specifically designed to work in Internet, we can calculate the diffusion possibilities of a web page of these characteristics.

5. Conclusions

Our main challenge, comparing this analysis with the analysis of most oral languages, is that research on this language is more recent and, furthermore, is presented here in new form: image. The acoustic form of oral languages as opposed to the image form of SL establishes not only linguistic differences, but also different ways of digital treatment.

Over time we have advanced in the linguistic knowledge of SSL and its digital treatment. It is therefore possible to achieve the main target of this Sign Library: to create a broad documental supply, essential for a language which has no accepted written register. This is why it is so important to offer a formal and standardised register, as the Sign Library offers in the videos in the other sections. This work, developed over the past three years, includes literary registers in its literature section and academic texts in the linguistic section on sign language, among others.

We know that this project, unique in Spain, is particularly followed by the deaf community, eager to discover how their ancestors signed or the origin of many of the signs they use today. This is all possible through the Internet, the ideal framework for persons who do not need to hear or to be heard in order to communicate.

Acknowledgements

This paper could not have been made without the support of Cervantes Virtual Library: www.cervantesvirtual.com and the economical support of Ministerio de Ciencia y Tecnología (Research Proyect nº BFF2002-010016).

References

A Basic Guide to ASL, U.S.A., URL: http://www.masterstech-home.com/ASLDict.html [Cited: 30/01/2004]

ASL Browser , [online], U.S.A., URL: http://commtechlab.msu.edu/sites/aslweb/browser.htm [Cited: 30/01/2004]

ASL University, [online], U.S.A., URL: http://www.lifeprint.com/asl101/index.htm [Cited: 30/01/2004]

Bellugi, U.; Klima, E. & Siple, P.(1975). Remembering in signs. In: Cognition 3: 2, - pp. 93-125.

Biblioteca Virtual Miguel de Cervantes, [On line], Alicante, Biblioteca Virtual Miguel de Cervantes. URL: http://www.cervantesvirtual.com/. [Cited: 30/01/2004]

Biblioteca Virtual Miguel de Cervantes, Biblioteca de Signos [On line], Alicante, Biblioteca Virtual Miguel de Cervantes. URL: http://www.cervantesvirtual.com/portal/signos. [Cited: 30/01/2004]

Biblioteca Virtual Miguel de Cervantes, Biblioteca de Signos [On line], Alicante, Biblioteca Virtual Miguel de Cervantes. URL: http://www.cervantesvirtual.com/portal/signos/autores.shtml. [Cited: 30/01/2004]

Biblioteca Virtual Miguel de Cervantes, Biblioteca de Signos [On line], Alicante, Biblioteca Virtual Miguel de Cervantes. URL: http://www.cervantesvirtual.com/portal/signos/cat_lingüística.shtml. [Cited:30/01/2004]

Biblioteca Virtual Miguel de Cervantes, Biblioteca de Signos [On line], Alicante, Biblioteca Virtual Miguel de Cervantes. URL: http://www.cervantesvirtual.com/portal/signos/cat_literatura.shtml.[Cited: 30/01/2004]

Bonvillian, J.D.; Folven, R.J. (1993): Sign language acquisition: Developmental aspects. In: Marschark, M. & Clark, M. D. (eds): Psychological perspectives on deafness. Hillsdale, NJ : Erlbaum, - pp. 229-265.

Fernández Villabrille, F. (1851). Diccionario usual de mímica y dactilología : útil a los maestros de sordo-mudos, a sus padres y a todas las personas que tengan que entrar en comunicación con ellos. Madrid, Imprenta del Colegio de Sordo-mudos y Ciegos.

Frishberg, N. (1975)Arbitrariness and iconicity: Historical change in ASL. In: Language 51: 3, - pp. 696-719.

Handspeak, [online], U.S.A., URL: http://www.handspeak.com [Cited: 30/01/2004]

Hervás y Panduro, L. (1795). Escuela española de sordomudos, o Arte para enseñarles a escribir y hablar el idioma español, dividida en dos tomos. [Tomo I]. Madrid, Imprenta Real.

Hervás y Panduro, L. (1795). Escuela española de sordomudos, o Arte para enseñarles a escribir y hablar el idioma español, dividida en dos tomos. [Tomo II]. Madrid, Imprenta Real.

Juncos, O. [Et al.] (1997) "Las primeras palabras en la Lengua de Signos Española. Estructura formal, semántica y contextual". Revista de Logopedia, Foniatría y Audiología , 17, , 3, pp. 170-181.

Pablo Bonet, J. (1620) Reducción de las letras y arte para enseñar a hablar a los mudos. Madrid, Francisco Abarca de Angulo.

Signhear ASL Dictionary, [online], U.S.A., URL:http://library.thinkquest.org/10202/asl_dictionary_text.html [Cited: 30/01/2004]

Silent Thunder Animations Page , [online], U.S.A., URL: http://www.velocity.net/~lrose/aniasl/anisign.htm [Cited: 30/01/2004]

Woodward, J. & Erting, C. J. (1975) Synchronic variation and historical change in ASL. In: Language Sciences. A world journal of the sciences of languages 37, - pp. 9-12

A Language via Two Others: Learning English through LIS

Elana Ochse

Dipartimento di Ricerca Sociale, Università del Piemonte Orientale
84 Via Cavour, 15100 Alessandria, Italy.
ochse@cisi.unito.it

Abstract

The complex intercultural activity of teaching/learning to read and write in a foreign language clearly involves a reciprocal cultural exchange. While trying to get students to efficiently learn the language in question, namely English, the teacher adapts to her pupils' culture and communication mode: in this case LIS or Italian Sign Language.

This paper attempts to demonstrate the complex process of developing a corpus for analysis of selected foreign language classroom exchanges. Here our emphasis is on face-to-face communication: what is imparted to the students by the teacher in Italian, how this information is transmitted or filtered by the LIS interpreter, what information the students eventually receive and how they react to it. A particular example of classroom activity has been filmed, transcribed and analysed from the points of view of successful communication, on the one hand, and failure or breakdown of exchange, on the other.

1. Introduction

A natural Sign Language, the dominant code in which face-to-face communication between Deaf people and other signers takes place, can be put on a par with an oral mode of communication (Yule, 1985; Ochse, 2004); however, in order to achieve literacy, the Deaf are obliged to learn another language with both a spoken and a written variant (usually the majority language of his/her area or country). Hence the "bilingual-bicultural" label which is often attached to Deaf signers (Swanwick, 1998; Prinz, 2002). Clearly the Deaf, who need a written language "to take part in the culture of the society in which they live" (Andersson, 1994) have a harder task than their hearing counterparts to learn the written language whose spoken equivalent they cannot hear. This may result in varying levels of second language literacy.

The subjects in our present study are Deaf Italian adults who have chosen to study English as a foreign language for personal interest and, if they are regular university students, to satisfy a credit requirement for their degree courses. A special project has been started for Deaf adults at the local university allowing them to follow experimental all-Deaf English classes with an emphasis on only written English (i.e. reading and writing) and assisted by a LIS interpreter.

From a certain point of view Italian and English are very similar since they have both a spoken and a written component. In the present situation Italian, the LIS-signer's second language, is likely to be the stronger written language because of more familiarity with it. On the other hand, English, like all foreign languages, is probably used only in classroom interactions and on some occasions in the external "linguistic landscape"[1].

[1] cf. Elana Shohamy: paper, entitled "Linguistic Landscapes, Multilingualism and Multiculturalism: A Jewish-Arab Comparative Study", presented at an international conference on Trilingualism, Tralee (Eire) on Sept. 4th 2003.

2. Method: data collection and presentation

In accordance with linguistic anthropological research methods (Duranti, 1997), a corpus of communicative events involving classroom discourse have been filmed. Meaningful excerpts from these ethnographic records (more than 25 hours of videotaped activity) have been selected and transcribed with the help of a native LIS (Lingua dei Segni Italiana) signer and linguistic expert.

In the present paper one of these excerpts, involving the teacher's communication in Italian (Column A), a translation of the latter into English (Column B), the interpreter's rendering into LIS or Italian of the teacher's or students' contributions (Column D) and the response or reactions of the class (Column C), has been analysed (See Table 1 below). A comparison between Columns A/B and D, i.e. the teacher's original or translated verbal communication followed by the LIS interpreter's rendering of the latter, can give evidence of success or failure in comprehension, language contact/interference and leakage.

As far as the transcription of the verbal and visual texts is concerned, for clarity we have opted for the simultaneous representation in four parallel columns of "utterances" or "speech events" instead of the "musical-score" format[2].

The lesson deals with the possessive form and, as is recommendable in Deaf didactics, has been enriched visually by projecting different slides on the screen.

The first slide portays a secretary in an office. Names, like *the secretary*, *Miss Smith* and *Mary* have previously been written on the board, in addition to various things that could be associated with her in the photograph (e.g. PC, laptop, portable computer, office, desk).

The second slide represents a woman holding a baby in her arms. Once again, different names, such

[2] Lucas (2002) quotes Ehlich ((1993) : " the musical-score allows the sequence of events to unfold from left to right on a horizontal line …" (44).

as *the baby*, *the mother* and *Joan*, have been written on the board.

The process is repeated with two more slides. Then the class are shown a few written examples of meaningful possessive phrases of the proper noun or common noun possessor + thing/person possessed (e.g. *Miss Smith's computer*, *Joan's baby*, *the baby's mother*).

3. Analysis

We have opted for an utterance/utterance analysis in the printed column format (column A vs D) to see if single communicative acts have been successful or not. The teacher explains that she has chosen a particular position so that she can point out things to the class on the screen. Then, to introduce the first possessor, she indicates *the secretary*, but feels the need to call on the class because she realises that their concentration is slipping. Before this interruption the interpreter has transmitted very little verbal information (GIRL), but probably sees the visual aid as an adequate alternative to a lengthy description. An image – the laptop computer – has attracted the students' attention and an animated signed conversation ensues. Since only one videocamera has been used, we have to follow the signing through the interpreter's words. Initially she tells the teacher that the lesson has been interrupted by the students' conversation, but then goes into the student mode, interpreting directly what different students are signing. One student is particularly enthusiastic about the laptop and reminisces about one with two other girls. But then she apologizes bi-modally to the teacher (sign + lip-pattern). She identifies the object as "everybody's dream".

When the students' conversation subsides, the teacher resumes her presentation and repeats who she was describing before the interruption ("the secretary"). To render the idea "This girl is a secretary", the interpreter concisely transmits the information in a question-answer form: GIRL*pl* – WORK – WHAT*q*, followed by the brief answer SECRETARY and the fingerspelling of the Italian equivalent. Showing adequate interest in the students' previous conversation the teacher makes reference to it and asks a question about the meaning of the acronym PC. In the interpreter's rendering she fingerspells C-O-M-P-I-U-T-E-R with an additional I, and then confuses the order of the letters PC, but quickly corrects herself.

Reference is then made to a number of phrases written as examples on the board (*the secretary's computer*, *Mary's desk*, *Miss Smith's office*) and containing different names for the same person (i.e. first name, common noun, title and surname). The interpreter "rewords" the message as follows: MESSAGE-MESSAGE-SAME-FIRST; EXAMPLE-SAME-BEFORE; WOMAN-NAME*pl*; NAME-SURNAME-NAME-SURNAME.

At this point a student goes back to the previous discussion about the laptop and asks if it also has a CD compartment.

In the second slide, where a child and a mother are introduced, a student reads the word *child* on the board phonetically. The interpreter fingerspells the word *bambino* but then signs and mouths BOY-GIRL to show the ambivalence of the English word *child*.

After *child* and *mother,* the interpreter feels the needs to list the following person nouns *Joan* (3rd) and *baby* (4th).

An interesting example of hybridisation occurs with GIOVANNA (common sign name for Giovanni followed by the fingerspelling A-N-N-A).

The choice of the word *bambino* by the teacher for both *child* and *baby* clearly confuses the students who ask for elucidation. The interpreter does not repeat this to the teacher but immediately starts explaining that *child* (bambino) can be either male or female. No mention is made of the word *baby*.

After this presentation of the two slides (containing examples of possessor and possessed), the teacher asks the class to write some sentences in their exercise books, one with a proper noun and one with a common noun.

She interprets "proper noun" as NAME-NAME-PERSON-MY and "common noun" as NAME-NAME-SAME.

A stretch of interpreting follows which corresponds to silence on the teacher's part.

4. Results

The following phenomena were found in the classroom interaction represented in Table 1:

a) Bimodal communication (sign + mouthing) of everyday utterances such as YES or SORRY.

b) Interpreter's initiative on two occasions, probably because she feared her previous interpretation had not been clear.

c) The use of facial expression, especially in questions like *q*READY, *q*WHAT, *q*UNDERSTAND, *q*ALSO CD-COMPUTER.

d) Indication of persons or things by gestures (*pl*) or gaze.

e) Particular LIS syntax in some questions or statements like READY*q*; MEANING-WHAT*q*; GIRL-WORK-WHAT*q*. SECRETARY.

f) Use of fingerspelling in which Italian words are spelt with the LIS alphabet, e.g. L-A, I-L, B-A-M-B-I-N-O.

g) Expression of plural form in LIS by repeating the sign with additional body posture, e.g. SENTENCE-SENTENCE; NAME-SURNAME-NAME-SURNAME; NAME-NAME.

h) Body posture and sign: portable computer (the action of carrying accompanies the laptop bag); abbreviation (short) for Personal Computer.

5. Conclusion

If the teaching had taken place directly in LIS, i.e. without the presence of the interpreter, we could have spoken of a single linguistic filter, but in this case the presence of Italian as everybody's common language created a double linguistic and cultural filter. This increased the risk of misinterpreting information and sometimes led to the understanding of different meanings from the ones that were intended.

A	B	C	D
mi metto davanti, così posso indicare le cose	I'LL STAND IN FRONT SO I CAN POINT OUT THINGS		(pl) MUST - STAND
abbiamo una ragazza che possiamo chiamare la segretaria	WE HAVE A GIRL WHO WE CAN CALL THE SECRETARY		+GAZE READYq - HAVE - GIRL
guardate Anna (………………….)	LOOK AT ANNA (……………………)	(students signing to one another) (invisible to camera)	(interrupts) GIRL (waves hands for attention) si, stanno parlando. Allora stanno … Si, in effetti, è molto bella questa foto col computer con la ragazza, dice Va ad Ar e An. Ti ricordi? E' bello. A … parlo del computer portatile molto carino. Scusa scusa (s+s). Stavo osservando. Giusto. Sogni di tutti. Vero. Sogno. Si, si (s+s)
è tuo sogno. Questa ragazza è la segretaria	IT'S YOUR DREAM. THIS GIRL IS THE SECRETARY		WELL ( ) – YOUR DREAM (pl) LIKE A LOT. GIRLpl WORK WHATq. SECRETARY. L-A S-E-G-R-E-T-A-R-I-A
e lei ha questa cosa che piace a Va: un computer oppure semplicemente con due lettere PC	AND SHE HAS THIS THING WHICH Va LIKES. A COMPUTER OR SIMPLY TWO LETTERS, PC.		THERE'S – LIKES – Va'S SIGN NAME – LIKES A LOT – PORTABLE COMPUTER _____ OR C-O-M-P-I-U-T-E-R (sic) OR PRONOUNCE SHORT C-P. NO P-C. SHORT P-C
qualcuno di voi sa cosa vuol dire questo PC? Si	DOES ANYONE KNOW WHAT PC MEANS? YES.	(students signing to one another)	MEANING WHATq Si, personal computer
e qua come vedete ho fatto una cosa simile. Ho chiamato la ragazza con questo nome e le ho dato anche un cognome e ho fatto vari esempi simile a quelli che abbiamo già fatto prima	AND HERE, AS YOU CAN SEE, I HAVE DONE SOMETHING SIMILAR. I HAVE CALLED THE GIRL WITH THIS NAME AND I HAVE ALSO GIVEN HER A SURNAME AND I HAVE MADE EXAMPLES LIKE THE ONES WE HAVE ALREADY MADE		pl SEE SENTENCE SENTENCE. SAME FIRST SAME WOMAN pl NAME pl NAME SURNAME NAME SURNAME. I GIVE SHOW EXAMPLE ALWAYS SAME BEFORE SAME
e ho fatto possedere anche altre cose come un ufficio, una scrivania	AND I MADE HER POSSESS OTHER THINGS SUCH AS AN OFFICE, A DESK		THEN - I - PUT (pl) OFFICE-DESK ______ THEN
		ALSO CD - COMPUTER	Anche cd del computer?

si	YES		YES - YES (emphatically)
child che vuol dire bambino o bambina (pointing at the word “child” on the board)	CHILD MEANING BOY OR GIRL	[kilɛd] ph	(pl) MEAN – CHILD – (fs) B-A-M-B-I-N-O - BOY-GIRL (s+s) – BOTH - SAME
poi mother si	THEN MOTHER YES	□ (student speaking) mamma	THEN (pl) MOTHER (pl)
poi Joan, è come Giovanna	THEN JOAN, LIKE GIOVANNA		3rd – J-O (fs) - SAME – GIOVANNI A-N-N-A (s+fs)
e poi baby che è un bambino (pointing at “the mother’s baby” on the board)	AND THEN BABY THAT IS A CHILD		4th – B-A-M-B-I-N-O (fs) – MEAN - SMALL CHILD(ph) - BABY
		UNDERSTAND – NOT – REPEAT □	
			□ MOM(pl) OWN – CHILD – MEAN – OWN – B-A-M-B-I-N-O (fs) OR-GIRL-BOY – BOTH – SAME (nods)
volete scrivermi almeno due frasi	WILL YOU WRITE AT LEAST TWO SENTENCES FOR ME		( ) NOW – PLEASE – YOU – YOU – MUST WRITE – TWO – SENTENCE - (pl) SENTENCE - WITH
con un nome proprio	ONE WITH A PROPER NOUN		NAME - PERSON - MY
e con un nome comune	AND ONE WITH A COMMON NOUN		(pl) SECOND - SENTENCE – PUT – NAME – NAME – SAME –
scrivetele sui vostri quaderni	WRITE IN YOUR EXERCISE BOOKS		YOU. OK ( )
si? avanti	OK? GO AHEAD.		TO YOU
adesso voi dovete scrivermi delle frasi nello stesso modo usando queste informazioni	NOW YOU MUST WRITE A FEW SENTENCES IN THE SAME WAY USING THIS INFORMATION		SO – YOU – MUST WRITE – SENTENCES – TWO – DIFFERENT – SENTENCE PROPER NOUN (s+s) (mouthing “comune”)
			PERSON (pl) – HAVE – POSSESS (pl)
			GIVE – WORDS – YOU – SEE – MEMORIZE – ELABORATE – BUILD – SENTENCE - SENTENCE
			BUT –NAMES - TWO – - DIFFERENT
			ONE – SENTENCE – NOUN - PERSONAL
potete per esempio	FOR EXAMPLE YOU CAN		SECOND (pl) –

dire “la mamma del bambino”	SAY: THE BABY’S MOTHER		SENTENCE – L-A (fs) MOTHER - OF (pl) - BABY
oppure il bambino della mamma o il bambino di Giovanna	OR THE MOTHER’S CHILD, OR JOAN’S CHILD		OR I-L (fs) BABY – HAVE – OF – GIOVANNI A-N-N-A
			MEANING USE – PROPER NOUN – COMMON NOUN
			UNDERSTOOD

Table 1: Transcription of a filmed extract of classroom interaction (University of Turin, 16th March 2002).

see	conventional orthography representing spoken Italian
SEE	translation into English of spoken Italian
SEE………….	English gloss of a LIS sign
I-SEE-YOU	single LIS sign glossed by more than one English word
S-A-W	when a word is fingerspelled, individual letters are separated by a hyphen
q	question
( )	pause (shorter than two seconds)
( )	off-topic, overlapping signing amongst learners
s	signing
(s+s)	lip-patterning and signing simultaneously
(s+fs)	signing followed by fingerspelling
ph	phonetic pronunciation
________	use of body posture
	overlap
+GAZE/-GAZE[3]	looking at or averting gaze from an addressee or object. Sometimes used as a form of placement.

Table 2: Transcription conventions (adapted and developed from Napier 2002)

[3] cf. Van Herreweghe (pp. 79-80).

6. Acknowledgements

This research is part of the Turin-based unit's contribution to the 2002 National Project (MUIR COFIN, project n. 2002104353) entitled "Strategies of Textual recasting for intercultural purposes". My special thanks to Claudio Baj (who helped me with the transcriptions from LIS to Italian), my Deaf students and Marco Gastini (for filming the classroom activities)

7. Bibliographical References

Ann J. (2001). Bilingualism and Language Contact. In C. Lucas (Ed.). *The Sociolinguistics of Sign Languages* (pp. 33-60). Cambridge: Cambridge University Press.

Duranti A. (1997). *Linguistic Anthropology*. Cambridge: Cambridge University Press.

Hamers J.F. (1998). Cognitive and Language Development of Bilingual Children. In I. Parasnis (Ed.). *Cultural and Language Diversity and the Deaf Experience* (pp. 51-75). Cambridge: Cambridge University Press

List G. (1990). Immediate Communication and Script: Reflections on Learning to Read and Write by the Deaf. In S. Prillwitz, T. Vollhaber (Eds.), *Sign Language Research and Application* (pp. 65-75). Hamburg: Signum.

Livingston S. (1997). *Rethinking the Education of Deaf Students*. Portsmouth: Heinemann.

Lucas C. (Ed.) (2002). *Turn-Taking, Fingerspelling and Contact in Signed Languages*. Washington D.C: Gallaudet University Press.

Marschark M., Lang H.G., Albertini J.A. (2002). *Educating Deaf Students: from research to practice*. Oxford: Oxford University Press.

Napier J. (2002). *Sign Language Interpreting: Linguistic Coping Strategies*. Coleford: Douglas McLean.

Ochse E. (2004) Language – English – Difficult// Question – You – Think – What? In M. Gotti, C.Candlin (Eds.), *TEXTUS XVII* (pp. 105-120). Tilgher: Genova.

Prinz P.M. (2002). Cross-Linguistic Perspectives on Sign Language and Literacy Development. In R.Schulmeister, H.Reinitzer (Eds.) (2002). *Progress in Sign Language Research*. In Honor of Siegmund Prillwitz (pp. 221-233). Hamburg: Signum.

Swanwick R. (1998). The teaching and learning of literacy within a sign bilingual approach. In S. Gregory, et al. (Eds.), *Issues in Deaf Education*. London: David Fulton Publishers Ltd.

Van Herreweghe M. (2002). Turn-taking mechanisms and active participation in meetings with Deaf and Hearing participants in Flanders. In *Turn-Taking, Fingerspelling and Contact in Signed Languages*. Washington D.C: Gallaudet University Press.

Yule G. (1985). *The study of language. An introduction*. Cambridge: Cambridge University Press.

Making Dictionaries of Technical Signs: from Paper and Glue through SW-DOS to SignBank

dr. philos. Ingvild Roald
Vestlandet Resource Centre
POB 6039 Postterminalen
NO-5897 BERGEN
Norway
E-mail: ingvild.roald@statped.no

Abstract

Teaching mathematics and physics in upper secondary school for the deaf since 1975, this author has felt the need to collect signs for the various concepts. In the beginning illustration of signs were pasted into a booklet. Then SignWriting appeared, and signs were hand-written and later typed into the booklet. With the 3.1 version of SignWriter, the dictionary program appeared, and several thematic dictionaries were made. With the new SignBank program, there are new opportunities, and I can fill in what I before just had to code. Last year a Fulbright research fellow and myself were collecting signs for mathematics, and these are transferred into a SignBank file. From that file various ways of sorting and analysing is possible. Here these various stages are presented, with a focus especially on the SignBank and the opportunities and limitations that are present in this program.

Paper and Glue

In Norway deaf students were educated according to the oralist method from 1880 until about 1960. Then signs were introduced in the schools trough the system called 'Correct Norwegian Sign Language' (C-NSL), a system supported by the Norwegian Deaf Association. Generally, deaf students were regarded as unable to grasp abstract ideas, and they were educated accordingly. Thus, when we started to question these old 'truths' about the abilities of the deaf students, and to offer education in more subjects and at more advanced levels, there were no signs in the language for the new concepts. Discussing this problem with the deaf people on the committee who 'proposed' or 'ran' the C-NSL, we agreed that the use of available signs from C-NSL, from other Scandinavian sign languages, from Gestuno or from American Sign Language, could be a basis for such technical signs. These were the sign languages for which I could get hold of dictionaries. (Dictionaries are listed in the references). All signs were discussed with my deaf students, and to preserve those signs I photocopied them and glued them into a booklet, according to theme. Examples are shown in figure 1. The process of choosing or creating the signs is described in an article published on the Internet (Roald 2000). Although the creation of signs should ideally be done by concensus in a population of native signers well versed in the topic for which the signs are to be used (Caccamise, Smith et al. 1981), this would not be possible in a population that is just entering a field of knowledge. The process that was used in our case, has been reviewed by Deaf teachers later on (Roald 2002), and was deemed appropriate. It is also a fact that new technical terms are coined in all languages when new needs arise,(Picht and Draskau 1985). It is also a fact that often a whole new vocabulary is made by one or a few persons who are working in a field and have authority. An example is Lavoisier (1743-94), who created the vocabulary of chemistry and laid down its rules of 'grammar' for times to come.

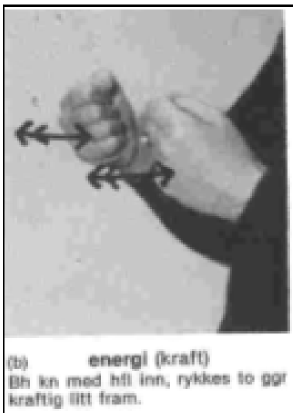
	**Energy** is the ability to perform work. Energy can have different forms, but the total amount of energy in the world is constant. In Norwegian, the term 'energy' may casually be used to mean 'force' (kraft) The term 'kraftkrevende industri' really means 'energy consuming industry' Electrical power (kraft) is really electrical energy
ENERGY Symbol: E	Energy is measured in units of J = Joule 1 J = 1 N·m = 1 W·s

Figure 1:'Energy' in paper and glue-version

SignWriter-DOS

With SignWriting came the possibility to *write* the signs rather than relaying on photos or drawings of persons signing. When I first met SignWriting, while visiting a Danish school for deaf pupils in 1982, I was fascinated and quickly adopted this way of preserving signs for my own benefit. With the arrival of the computer program, the SignWriter, came the opportunity to type and preserve the signs in a neat way. In its first versions the SignWriter did not have a dictionary program, but by utilising the word-processing possibilities and the possibilities to write in the Roman alphabet as well as in sign symbols, it was nevertheless possible to make short lists of signs for specific themes. An example is shown in figure 2.

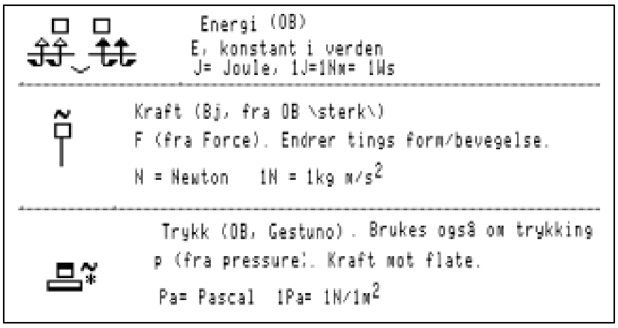


Figure 2: Energy etc. in early SignWriter version

SignWriter®-DOS Dictionary Program

In the early to middle 1990's I was given the task by my resource centre to develop materials for the teaching of deaf students. Signing was by now well established as the language of the deaf education in Norway, even if Norwegian still was necessary for dealing with the outside world. With the 3.1 version of the SignWriter program came the attached Dictionary program. This program made it possible to create real dictionaries of signs for concepts and words. Each sign was written separately either in the dictionary itself, or uploaded from a written sign text. Each sign had to have a name, and the dictionary was sorted alphabetically by these names. Sometimes more than one sign would correspond to the same Norwegian word. They might be variant signs for the same concept, or they might be signs for different concepts covered by the same name-word. These were coded by (1), (2), (3), etc. Often a short explanation would also go into the name field. The source of the sign would also be in the field, as a coding or a short note. As the writings, or spellings, of signs are not yet established, at least not in Norwegian Sign Language, I often gave multiple written versions of the same sign. These were coded by using (i), (ii), (iii) etc. Examples are given in figure 3.

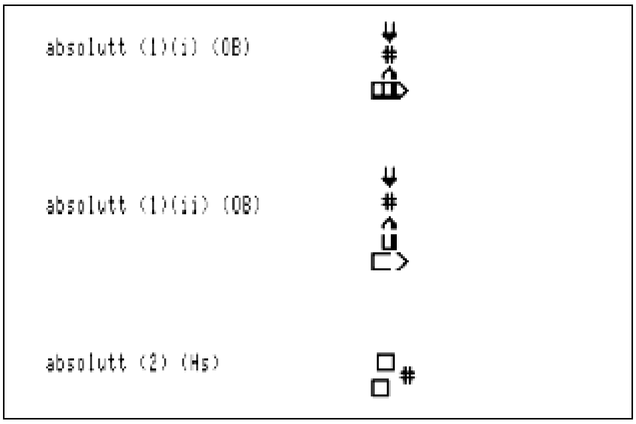


Figure 3: 'Absolutely' in SW 4.3 Dictionary program

Problems with SW's Dictionary Program

The Dictionary program has a feature called 'merging'. My hope was to use this feature to build a large dictionary from several smaller dictionaries. That way it would be possible to make a dictionary of 'Spring Flowers' and merge this with a similar dictionary of 'Wild Flowers', one of 'Summer Flowers' and so on, and merge these together and make a dictionary of 'Flowers', which again could be merged with other small ones to make 'Botany' and 'Biology' and 'Science' and finally 'Norwegian Sign Language'. Several attempts along this road were made, but the program would often fail in the merging process. Figure 4 shows result of a failed attempt to merge two smaller dictionaries. This was a setback, as it is considerably harder first to make a large dictionary and the weed out everything that is not inside your chosen theme. Each time a new sign is added, it has to be added seperately to each of the appropriate theme dictionaries, rather than doing the merger process over at regular intervals. The building of dictionaries for signs other than my main subject, physics, therefore halted. In my computer are several small dictionaries for marine life, for instance. They will be used to fill the Norwegian SignBank.

Another time-consuming problem, not related to the Dictionary program, has been the changes in symbols and which key they are allotted to. The newer versions have not been compatible with the older ones, and strange pictures have resulted from this.

In addition, SignWriter and the Dictionary program can not be run from newer computer platforms, such as Windows 2000 or Windows XP.

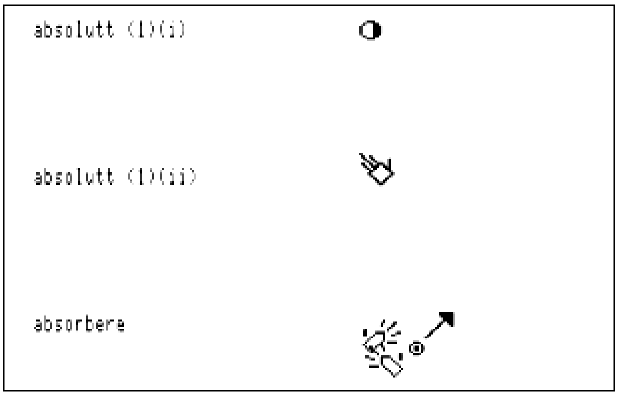


Figure 4: Result of merger failure with SW Dict program. The sign called 'absorb' is really 'language'

The SignBank® Program

The SignBank is a very different program. It is built as relational databases in a FileMaker® environment. The program is not suited to write signed texts, but is a toolbox for creating good Sign dictionaries in a variety of styles. It can be searched alphabetically by sign-names, but it can also be searched from the sign itself. In addition, it can be searched by themes or other criteria that the editor may choose to use. Chosen signs, making up a large or small dictionary, can be printed in a variety of ways. Video explanations of the signs can be added, as can photographs or drawings like those from the paper-an-glue era. Other illustrations and explanations, in spoken-language text or in signed-language text can be added.

The recording of signs into the SignBank is rather time consuming, but this is necessary to make the program able to sort by sign. To make this sorting by sign possible, a standard sequence of the different symbols in SignWriter, along with a few extras from the Sutton MovementWriting® system, is established. The full version, called SSS-04 (Sign Symbol Sequence) a large number of symbols, a shorter and more compact version created for the American Sign Language contains a smaller number of symbols in 10 categories, divided into groups, again divided into symbols and their variations. Now, the large numbers may seem overwhelming, but the system itself is largely transparent, and the huge number stems from the fact that a hand shape can have 16 rotations and 4 'flippings' (shadings) for the right hand alone, making a total 128 symbols for that handshape (both hands included), all neatly derived from one basic symbol. With all possible hand shapes from all the signed languages, this makes the inventory of hand symbols huge. In addition, there are the symbols for movement (again having variations in direction and size, as well as which hand is moving) and dynamics, and the symbols for the face and other parts of the body, both as articulators and as places of articulation. The symbol sequence is used both in the SignWriter programs and in the SignBank, and will constitute the International Phonetic Alphabet. In the SignBank, this Symbol Sequence is used to order and look up the signs by the signs themselves. Thus, it becomes possible to have a dictionary for one signed language only, with definitions and explanations in the same signed language, without having to restore on any spoken language. It will also be possible to have dictionaries between two or more .signed languages.

For a sign to be recorded, it first has to be written in SW-DOS or SW-JAVA, and then made into a .gif-file or a .png-file. The writing rules of signs in SignWriting are still somewhat ambiguous, as writing rules for the different signed languages have not had time to make the orthography settle. Thus, a writer may have several ways of writing the same sign, as a way to for the signing community to settle for one or the other. This I have done by using the coding (i), (ii), (iii), etc., for different ways of writing the same sign.

The term 'spelling' in SignBank parlance, means the symbols chosen for the ordering of the sign into the sequence, and the ordering of these symbols. For that purpose, most signs are seen as consisting of three 'syllables': starting configuration, movement, and ending configuration. The rules now are:

1. Initial dominant hand in shape and rotation and shading
2. Non-dominant hand similarly, if that hand is taking part in the sign.
3. Initial symbol for other articulators
4. Place of articulation
5. Movement of dominant hand (fingers first, then hand, etc.)
6. Movement of non-dominant hand
7. Movement of other articulators (brows, eyes, mouth, …)
8. End dominant hand
9. End non-dominant hand
10. Dynamics

All these steps are optional, except for step 1. A few signs will have only this one symbol: most of the letters and numbers are given that way. Also, the few non-manual signs will have step 3 only.

For step 4, the place of articulation, extra symbols may be required that are not written in the sign. For the written sign, the placement in relation to the body is given by the structure of the written sign and its relations to the guiding (imagined) lines in the text. These are not part of the spelling for entering the sign into the SignBank. For use whenever necessary, symbols depicting the body and place of articulation are part of the SSS.

Once all the relevant symbols are entered into the spelling of the sign, the sign should be saved into the bank. A copy of the sign and the spelling can be made, for use with other signs that share the same features. Sometimes the exact same sign will cover more than one word in the spoken language (as for any two-language dictionary). Editing can also be done on this copy, so that a sign varying from the first in one or a few of the symbols can be entered with less entering work.

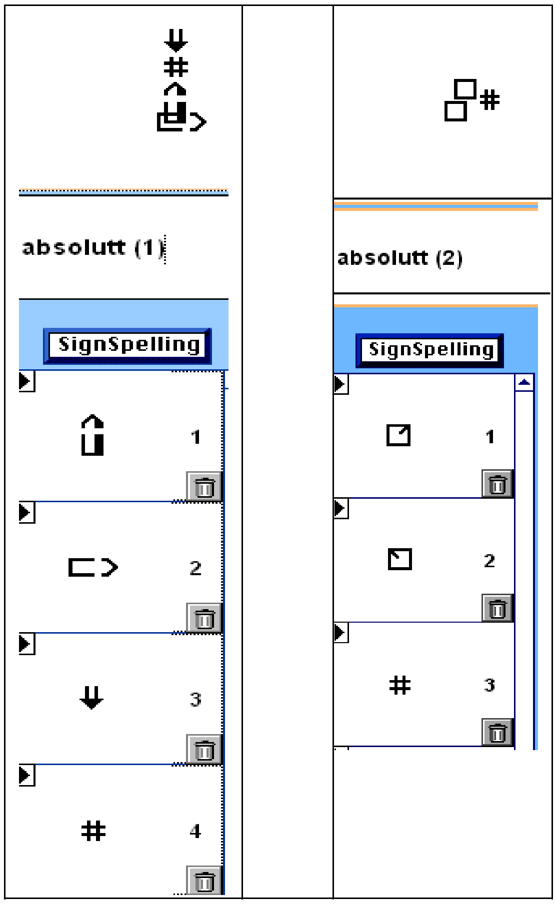


Figure 5: Sign spellings of 'absolutely'

In addition to entering the spelling of the sign, and a word-name for it, the editor has the possibility of entering linguistic data, the sign in a sign-context or a sign-text explanation of the sign like in any one-language dictionary. It is also possible to make dictionaries covering more than one signed or spoken language. Video, animations and still picture illustrations may also be added. In all, the possibilities are available for whatever one may want, as the program itself may be augmented with new features in new related database files. Below is shown a few of the features in the program

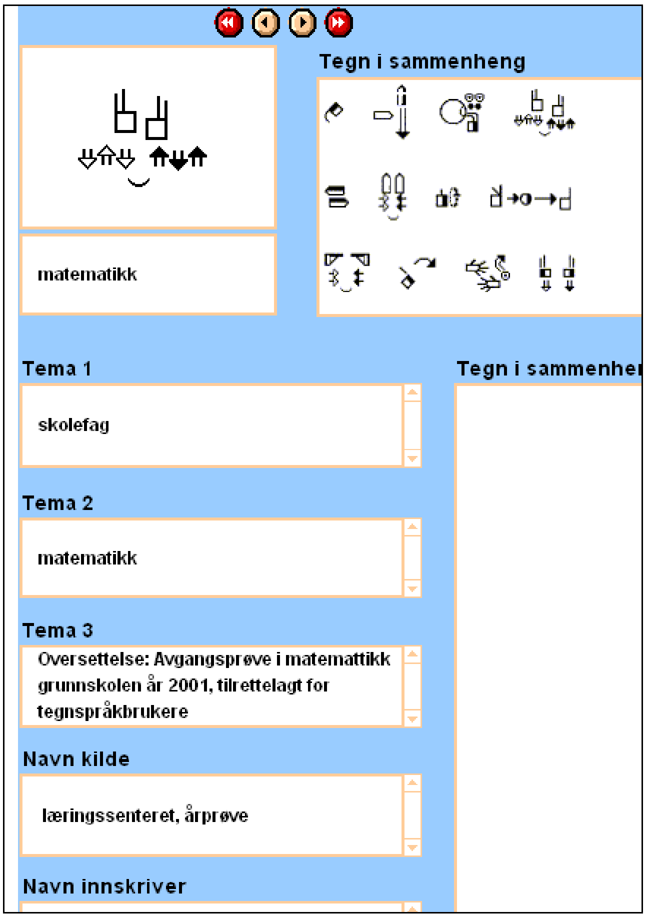


Figure 6: Linguistics page for 'mathematics'

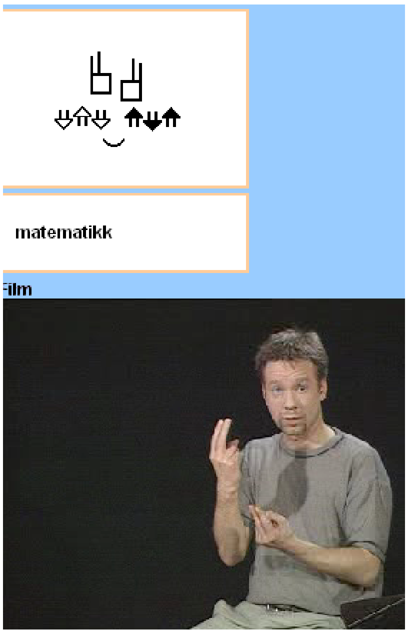


Figure 7: Page with snapshot from video

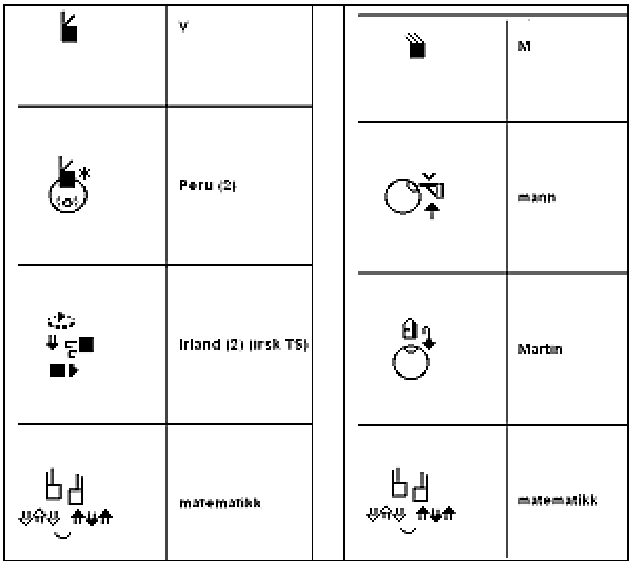


Figure 8: Parts of dictionary, sorted by sign or word

In a short article like this one, it is not possible to present all the features of a program like the SignBank. Suffice it to say that the program opens a new era in dictionary creation for the signed languages, combining the written sign, sign illustration, video of sign and of sign in context, translation between signed languages and between signed and spoken languages. We have only begun to scratch these possibilities.

References

Author unknown (1975). GESTUNO International Sign Language of the Deaf; Language Gestuel International des Sourdes. Carlisle, UK, The British Deaf Association.

Caccamise, F. and H. G. Lang (1996). Signs for Science and Mathematics: A resource book for Teachers and Students. Rochester, NY, Rochester Institute of technology.

Caccamise, F., N. Smith, et al. (1981). "Sign Language Instructional Materials for Speech, Language and Hearing Professionals." Journal of Academy of Rehabilitative Audiology **1981**(14): 33-61.

Norsk/nordisk tegnspråkutvalg (1976). Nordisk Tegnspråk, 155 tegn vedtatt på Nordisk tegnspråk-seminar i Malmø 1975. Bergen, Norway, Norske Døves Landsforbund

Picht, H. and J. Draskau (1985). Terminology : an introduction. Guildford, University of Surrey Department of Linguistic and International Studies.

Roald, I. (2000). Terminology in the Making - Physics terminology in Norwegian Sign Language, The Deaf Action Committee for SignWriting. URL :hhtp://www. SignWriting.org/forums/lingustic/ling032.htm

Roald, I. (2002). "Reflection by Norwegian Deaf Teachers on Their Science Education: Implications for Instructon." Journal of Deaf Studies and Deaf Education **7**(1): 57-73.

Tegnspråksutvalget (1988). Norsk Tegnordbok (=Norwegian Sign Dictionary). Bergen, Norway, Døves Forlag.

Searching SignWriting Signs

Steven Aerts*, Bart Braem*,
Katrien Van Mulders†, Kristof De Weerdt†

*University of Antwerp
Campus Middelheim, Building G
Middelheimlaan 1
B 2020, Antwerp, Belgium
{bart.braem, steven.aerts}@ua.ac.be

†University of Ghent
Rozier 44
B 9000 Gent, Belgium
{katrien.vanmulders, kristof.deweerdt}@ugent.be

Abstract

At the moment the publication of the first written Flemish Sign Language (VGT) Dictionary is in progress. It consists of VGT glossaries and allows its users to lookup the signs for over 2000 Dutch words. The signs are written in SignWriting.
We have established an electronic representation of this sign language dictionary. Searching for signs starting from a Dutch word works straightforward. The opposite, receiving results ordered by relevance, has never been develloped before. In this paper we explain how we have worked out such a system.

1. Introduction

We have developed an online database driven dictionary system currently containing about 5000 signs(Aerts et al., 2003). These signs were (and are) collected by researchers of the university of Ghent and written down in SignWriting using SignWriter DOS(Gleaves,). Our system can convert these binary files to SWML, which is an XML-based SignWriting representation language(da Rocha Costa and Dimuro, 2003). The major advantage of SignWriting in general and SWML in particular is that it is very lightweight and thus ideal for the web. Our database is modelled on the SWML structure to contain exactly the same information.
The SignWriting system itself is a practical visual writing system for deaf sign languages, composed of a set of intuitive graphical-schematic symbols and simple rules for combining them to represent signs(da Rocha Costa and Dimuro, 2003). It was invented by Valerie Sutton inspired by her already created choreographical writing, called language DanceWritingSutton, a) Sutton, b). SignWriting symbols represent the body parts involved and the movements and face-expressions made when producing signs. This way in an electronic representation each symbol is stored along with its transformations: translation, mirror and rotation. In this paper we provide an outline to construct an intuitive, user-friendly, yet powerful search by sign system for SignWriting. All information can be extracted from the signs only, without external information e.g. the position of the hand or the dominant hand.

2. The manual approach

Searching using only the SignWriting signs has never been done before. Searching for the meaning of a sign in a database containing filmed signs manually enriched with extra semantic information, however, is common practice. It usually consists of selecting the type and direction of the movement, the location on the body where the sign is made and finally the hand form. This information needs to be added to each individual sign, causing a big slowdown in dictionary development due to tedious manual work. When creating a huge, constantly evolving dictionary this is highly undesirable.

3. Semantic view on SignWriting

The first question we should ask ourselves is whether the following common search information can be extracted from SignWriting signs:

- Type of movement: the movements are pretty well described as they are represented by different symbols, but whereas it is rather easy for human beings to find the only physiologically feasible possibility, it is difficult to find the matching moving body part with a computer.
- Direction of the movement: this is almost impossible to extract. Because of the two-dimensional representation in SignWriting the difference between horizontal and vertical movements is - again - only easily detectable by human beings.
- Location: a course grained distinction between zones of positions should be possible, but only when a body part is touched. When no body part is touched, the location can only be extracted from the SignWriting symbols by considering the most likely physiological positions.
- Hand form: is very accurately defined in SignWriting through the use of different symbols. Hand forms will obviously be the key feature to search by.

4. Elaboration

4.1. User input

The user first specifies the hand form. Then he specifies which body zones are touched (head, torso or legs) and the way in which they are touched (touch, grasp, in-between, strike, brush contact or rub contact).
It is also possible to specify the orientation of the hands (palm up, back of the hand up or side view). This does not change the essence of the search.

4.2. Processing

Selecting all signs that include the specified hand form(s) is obvious. The different types of touching are also very well depicted in SignWriting, but one type of touching has multiple variations. The body zones involved and the type of contacts are accurately specified. If multiple zones or contacts are involved, matching them is difficult.
It is however possible to parametrise the goodness of a match. The goodness-measure we would like to use is the product of the distances between the touch and the middle of the corresponding body zone. When measuring the goodness from a match of n contacts c_i $(i = 1..n)$ and the corresponding body zones z_i this results in the following function: $\mu(c_1, ..., c_n, z_1, ..., z_n) = \Pi_{i=1}^{n}((c_{ix} - z_{ix})^2 + (c_{iy} - z_{iy})^2)$

4.3. Results

We are able order matches, using this measure: the closer the match the lower the goodness-measure. Dropping very bad matches, which have a very high goodness-measure, is also possible. The ordering does not happen in a natural way (Butler, 2001; Sutton, 2004), because for that to be possible manually added information about e.g. the dominant hand would be necessary.

5. Issues

5.1. Precision & Recall

Precision is described as *the ratio between the relevant retrieved signs and the retrieved signs*, whereas recall stands for *the ratio between the relevant retrieved signs and the relevant signs*.
In this case: is the user able to clearly specify what he wants? If this is problematic, improving usability will be possible by broadening the search to closely matching symbols.

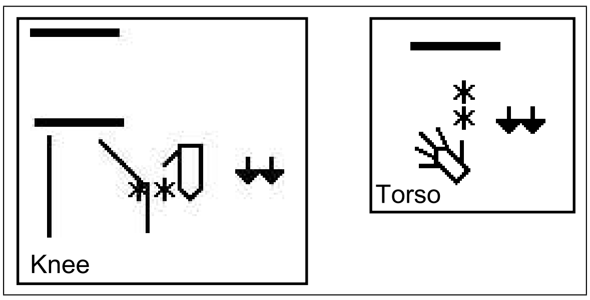


Figure 1: Determining the body zones

5.2. Performance

Selecting signs through the right symbols and contacts is a database issue where performance is not at stake. A well designed database will process those queries fast enough. Determining the body zones is rather straightforward and can also be done in the database. If we see one thick black horizontal line, for example, we know the chest is meant, whereas two lines depict the legs and hips. Figure 1 illustrates this principle.
Calculating the goodness-measure will be done over a very limited number of matching body-symbols and contacts: a sign containing four contacts is extremely rare. The number of comparisons will be low and will not affect the global performance by one order-of-magnitude.

6. Future work

We do not have an implementation of the search algorithm right now, which is the only missing part in our dictionary system. We expect to have a working system ready soon, because the method we have described is pretty straightforward to implement.
The structure of our system is built with the algorithm in mind. The exact mapping of SWML on the relational database prevents loss of information about the signs. Every single symbol can be traced back to its containing sign, allowing fast lookups of relevant signs.

7. Conclusion

The great advantage over existing systems is the fact that all information originates from the signs only.
This system is intuitive for a user with basic SignWriting knowledge. Its friendliness will largely depend on the used interface but can be improved with the goodness measure and a broadened search. The real strength of our system lies in the use of the very well specified SignWriting hand forms, which compensates for the vague movements.
Because of the use of databases, SWML and relatively simple calculations, this method is also straightforward to implement. Most important, the Deaf Community and its researchers will benefit by this new search method since it allows for easier dictionary-searching. Moreover the system can be used as an online reference for the meaning of SignWriting signs.

8. References

Aerts, Steven, Bart Braem, Kristof De Weerdt, and Katrien Van Mulders, 2003. `http://plantijn.ruca.ua.ac.be/˜aertsbraem/interface/` - Vlaams Gebarenwoordenboek.

Butler, Charles, 2001. An ordering system for SignWriting. *The SignWriting Journal*, 0(1).

da Rocha Costa, Antonio Carlos and Gracaliz Pereira Dimuro, 2003. SignWriting and SWML: Paving the way to sign language processing.

Gleaves, Richard. `http://signwriting.org/forums/software/sw44/` - signwriter dos computer program.

Sutton, Valerie, a. `http://www.dancwriting.org/` - DanceWriting: Read and write dance.

Sutton, Valerie, b. `http://www.signwriting.org/` - SignWriting: Read, write, type sign languages.

Sutton, Valerie, 2004. Sutton's SignSpelling guidelines 2004. SignWriting Library Database, DAC@signwriting.org.

Sign Printing System - SignPS

Inge Zwitserlood
Viataal
Theerestraat 42, NL-5271 GD Sint Michielsgestel, The Netherlands
i.zwitserlood@viataal.nl

Doeko Hekstra
Handicom
Oranjelaan 29, NL-3843 AA Harderwijk, The Netherlands
info@handicom.nl

Abstract

The development of the Sign Printing System (SignPS) is based on the need of a way for sign language users and teachers to compose pictures of signs without having considerable drawing skills, to store these pictures into a database and to retrieve them at wish for several purposes. The sign pictures are abstract but nevertheless recognizeable without specific training. The programme is not developed for scientific purposes, but for use by the general (signing) public.

1. 'Drawing' sign pictures

Similar to languages, sign languages often require a static representation, that can be used in print and processed in one's own pace. Currently, several types of static representations: photographs, drawings, glosses and several notation systems are used to represent single signs, sometimes also for the representation of sign sequences. The reason for yet another system is that the existing systems have several disadvantages, that are overcome by Sign PS

Glosses, first have the disadvantage of not giving information on the shape of a sign. Second, since glosses are labels taken from spoken languages, whose grammatical structure is often considerably different from the sign language, much of the information that is present in a sign cannot be expressed by words or affixes of these spoken languages. Various subscripts and superscripts are then needed to represent this information.

Disadvantages of photographs and many drawings are the unnecessary details they show (clothes, hairstyles), that can distract the onlooker from the message. Photographs show particular persons, drawings have particular styles. As a result it is seldom possible to use separate photographs and drawings to construct coherent representations of sign strings. Furthermore, most people's drawing skills are not sufficient to make drawings of signs, and photographs require special equipment and additional adaptations in order to represent the dynamic part of signs (such as the movement of the hands).

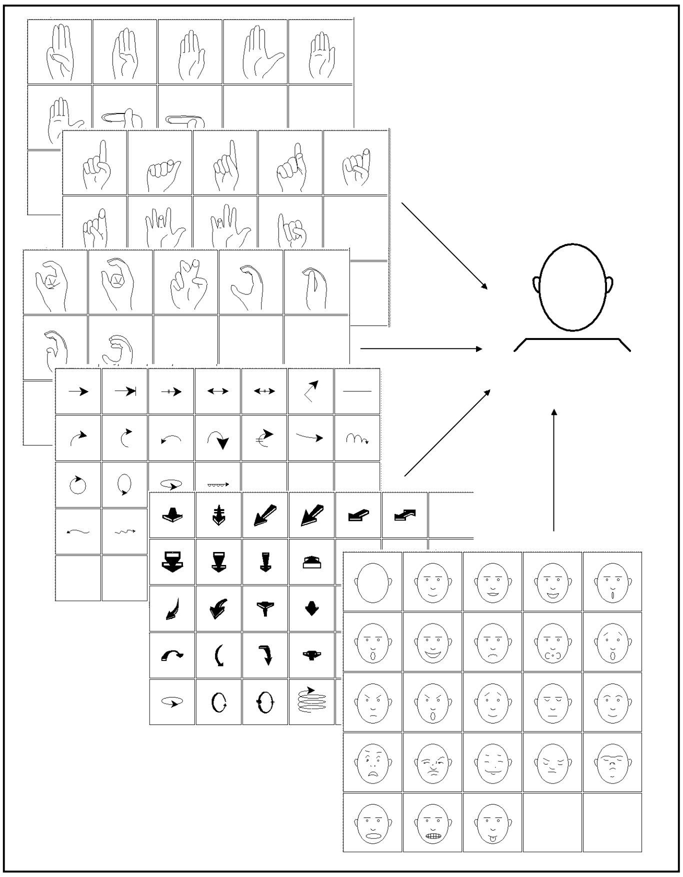


Figure 1 Components from restricted sets can be added to a sign picture.

Thirdly, although most notation systems (e.g. SignWriting, HamNoSys, KOMVA) do not entail these problems, they are not userfriendly for common language users in general, because special training is needed to learn to use them, and, more importantly, in some groups of sign language users there is a general resistance against the use of such systems for common use.

The Sign Printing System overcomes these problems by offering everyone with basic sign language skills a tool for quick and easy construction of sign pictures. The program opens with the contours of the head and shoulders of a signer. Handshapes can be chosen from a limited set of handshapes and added to the picture. These can be moved, copied, rotated or mirrored into the desired orientation. In the same way, arrows and other movement symbols can be chosen from limited sets, added to the picture and edited. Furthermore, particular facial expressions are composed by choosing and/or editing face components: eyes, eyebrows, mouth and nose. Subsets of these sign components are shown in Figure 1.

3-Dimensionality is suggested by 3-dimensional movement block arrows for movements towards and away from the signer and by varying the size of the hands. A large-sized hand gives the impression that it is closer to the onlooker than a small-sized one, as illustrated in Figure 2.

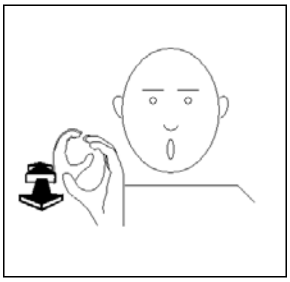


Figure 2 3-dimensionality in a 2-dimensional picture (NGT sign for ‘tube on shoulder’)

These sign pictures are rather abstract in that they only contain the minimal number of components necessary for understanding the sign. Because of this abstraction, they can also be easily combined to form sign strings. On the other hand, the abstraction in the sign pictures is not so extreme that special training is required for learning to recognize the signs.

At present, the Sign Printing System contains the sign components needed for signs from Sign Language of the Netherlands (henceforth: NGT). It will be fairly easy for the developers to make adaptations for other sign languages (such as different sets of handshapes).

2. Storage and retrieval

The Sign Printing System has in common with photographs and drawings that the sign pictures are stored as whole units. Once stored, a user can retrieve the sign pictures as whole units and does not need to compose a particular sign picture anew every time it is needed.

An innovative part of the Sign Printing System is the database. A sign picture that is stored in the database must be connected to a concept and to a gloss. In case of synonymic signs, it is possible to connect more than one sign picture to one concept (and one gloss). For instance, NGT has several synonymic signs meaning ‘good’ that can all be stored with the same gloss. This facilitates retrieval, since synonyms will not be overlooked. A sign picture can also be connected to more than one gloss (depending on the language from which the gloss stems). For instance, NGT has only one sign meaning ‘cat’, whereas Dutch has two words with that meaning, *kat* and *poes*. The sign picture can be thus labelled with both glosses, but still be stored as one picture. A number of concepts are present in the database from the start. They are ordered in a semantic hierarchy. Although it is not possible to change the database structure, a user can add concepts and glosses to the database and even add categories to the semantic structure.

Retrieval of sign pictures is possible in three ways. First, the database can be searched by gloss name, which is a common way for retrieval in many sign databases. Second, a user can search for sign pictures within the hierarchically structured semantic fields in the database. By choosing a particular semantic field, the user is shown the subset of the gloss names of the signs that are in this field in the database. This is illustrated in Figure 3.

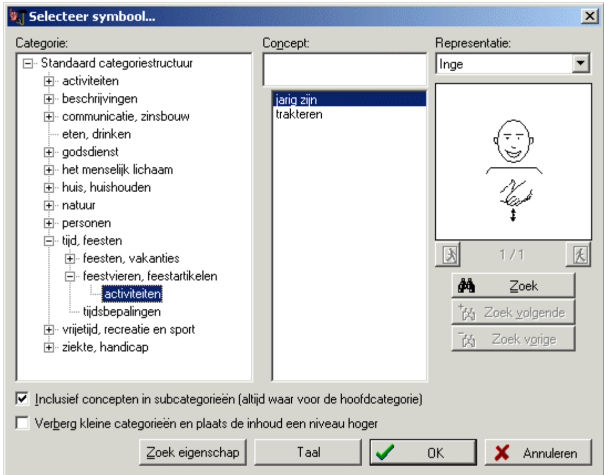


Figure 3 Searching within the semantic field of ‘celebration activities’

Third, a user can search sign pictures by selecting components of the signs, viz. handshape(s) and/or place of articulation. For instance, in Figure 4 the results are shown of a search operation for a sign with a particular handshape that is made in front of the chest. The particular orientations of handshapes in signs is not taken into account in the handshape search facility.)

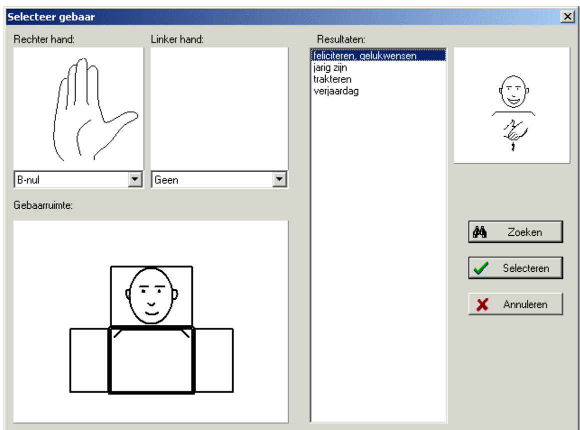


Figure 4 Searching by sign components

3. Use of the sign pictures

The Sign Printing System is part of a range of software applications using communication symbols and databases (sharing the same format) holding these symbols (such as Bliss, PCS and Picture this), called Symbol for Windows™. Among the applications are a plain word processor, an email program and several educational tools. The sign pictures stored in the database of the Sign Printing System can be used directly in these applications. Retrieval of these pictures is fast, and additionally, other elements (pictures, photographs or symbols) can be retrieved from the connected databases and used in the same application. For instance, one can combine the picture of the sign for 'cat' with a picture of a cat, and/or with the Dutch word *kat*. The combined use of sign pictures and a photograph is illustrated in an item of a multiple choice test in Figure 5.

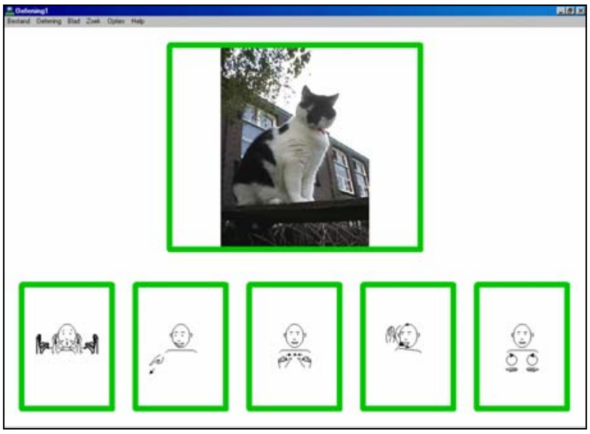


Figure 5 Multiple choice test for NGT idiom with SignPS pictures

Common Windows programs do not have direct access to the databases. The sign pictures can be retrieved from the database and stored as graphic files with an export tool that is included in Symbol for Windows. These pictures can easily be inserted in applications such as Word or Powerpoint. The sign pictures can also be cut and pasted into these applications. An example of a lecture using an NGT sentence is shown in Figure 6.

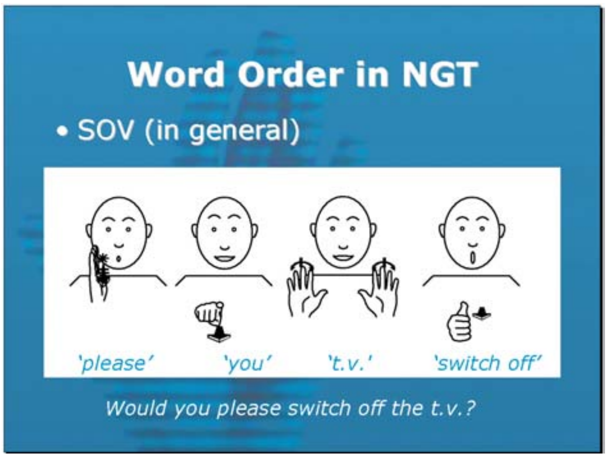


Figure 6 Powerpoint presentation with SignPS pictures

4. Further developments

The Sign Printing System is still under development. It has not yet been used and tested by large user groups. A first working version is distributed and a pilot course in the use of the program has recently been taught to a small user group (NGT teachers and speech therapists at Viataal). This group currently evaluates the program and their first reactions are very positive. A preliminary inventory of desirable adaptations shows that the set of viewing angles of handshapes should be extended and that the user-friendliness needs to be slightly improved.

A new Gesture Representation for Sign language analysis

Boris Lenseigne*, **Frédérick Gianni***, **Patrice Dalle***

*IRIT
Université Paul Sabatier 118 route de Narbonne 31062 Toulouse cedex 4
{lenseign,gianni,dalle}@irit.fr

Abstract

Computer aided human gesture analysis requires a model of gestures, and an acquisition system which builds the representation of the gesture according to this model. In the specific case of computer Vision, those representations are mostly based on primitives described from a perceptual point of view, but some recent issues in Sign language studies propose to use a proprioceptive description of gestures for signs analysis. As it helps to deal with ambiguities in monocular posture reconstruction too, we propose a new representation of the gestures based on angular values of the arm joints based on a single-camera computer vision algorithm.

1. Gesture representation

1.1. Previous work

Most of the descriptions of Sign language vocabulary relies on linguistic studies and are those used in notation or transcription systems, SignWriting[1], Hamnosys[2] (Prillwitz and al., 1989). In the case of computer aided Sign language analysis, we distinguish systems using a specific hardware such as data gloves (Braffort, 1996)(Starner T., 1998)(Vogler C., 1999) and those using cameras. Data gloves based applications process directly on values provided by the sensors. in he case of a computer vision system, gesture model cans be bidimentionnal or tridimentionnal. When several cameras are used (Wren C., 1999)(Vogler C., 1998)(Somers M.G., 2003) $3D$ reconstruction is possible and gestures can be analyzed directly in $3D$ space. In the case of a single camera, gesture analysis can be performed directly in $2D$ images (Starner T., 1998)(Tanibata N., 2002) or some additional image processing has to be performed for a partial $3D$ estimation. In this case, visual aspect of gestures is deduced from $3D$ models (Horain P., 2002)(Athitsos V., 2004), or $3D$ model is used to constrain the reconstruction (Lenseigne B., 2004). Both solutions leading to ambiguities.

Thus notation systems and vision-based gesture analysis systems use a representation of signs derived from a tridimentionnal perceptive description. Gestures are located in a speaker-centered frame (ScF) (fig. 5) but described from external point of view. Those descriptions are based on the definition of a set of elementary motions and for each elementary motion a set of parameters. Widely used primitives are, straight, curve and complex motions. Such a classification is only suitable for standard vocabulary description, leads to a classification of gestures in terms of geometrical primitives and to a description of gestures from the observer's point of view.

1.2. A new gesture representation

A different way to represent gesture is to use a proprioceptive point of view. In such a case, motion analysis and classification rely on the way gesture is performed. This approach is presented in recent linguistic research (Boutet, 2001) which suggests that an articulation-based representation may have appropriate properties to allow the representation of the function of the gesture. So that, using joint values to represent gesture is an interesting choice. This assumption leads us to propose a method, based on a single camera to compute a gesture representation based on joint angle evolution in time.

2. Computing articulation values from a single image

Articulations values calculation is performed in two stages : a geometrical reconstruction of the $3D$ posture of the arm and the computation of corresponding articulations values. As we use a single camera, a direct $3D$ reconstruction of the arm is not possible, and the geometrical method provides us with a set of four possible configuration of the arm in a given image. A configuration is represented by the $3D$ Cartesian coordinates of each joint (shoulder, elbow and wrist). Those coordinates are grouped together to form a set of four possible motions for the arm and joint values can be computed for each trajectory to build articulation-based motion representation.

2.1. Geometric resolution

In this section we describe how to reconstruct a set of possible $3D$ pose of a human arm from a single picture using a single calibrated camera. This configuration is defined by the position of each segment's limits (shoulder, elbow and wrist) in Cartesian coordinates. Given an image, we are able to reduce the space of possible poses for the arm to four configurations only using a simple model of the scene, the camera and some assumption about it.

2.1.1. Requirements

Our technique is based on several assumptions, which may be crippling under an uncontroled environment. However they could be raised if reconstruction can be performed with a scale factor which does not affect joint values computation.

[1] `http://www.signwriting.org/`

[2] `http://www.sign-lang.uni-hamburg.de /Projects/HamNoSys.html`

Acquisition device : The acquisition device is made up of a single camera, which has been calibrated in order to be able to calculate the equation of the projective ray across a given pixel, which suppose that the perspective transformation matrix $\mathbf{C}$ is known. Many techniques of calibration were previously proposed, for instance in (Gurdjos P., 2002) or (Heikkilä, 2000).

Tracking the articulations : We also make the assumption that we are able to identify the $2D$ positions of the three articulations of the arm in the image. Tracking techniques abound and depend on the problem to solve (degree of homogeneity of the background, the use of markers, motion models, etc...). The precision needed for tracking depends on the precision needed for reconstruction. A study of the influence of tracking errors on reconstruction can be found in (Lenseigne B., 2004).

Arm pose : We only speak here about rebuilding the posture of an arm, without considering the hand. Within the framework of a geometric resolution, we define the posture by the position in space of the articulations (shoulder, elbow, wrist), i.e. if coordinates are expressed in the camera frame :

- for the shoulder : $P_1 = (X_1, Y_1, Z_1)^T$
- for the elbow : $P_2 = (X_2, Y_2, Z_2)^T$
- for the wrist : $P_3 = (X_3, Y_3, Z_3)^T$

Using this representation, the estimating of the posture of the arm is reduced to the calculation of three points in space.

2.1.2. Geometrical model of the arm

The arm is modeled by the articular system connecting the shoulder to the wrist. This system consists of articulations (a ball-and-socket joint for the shoulder and a revolving joint for the elbow) connecting rigid segments (arm and forearm) noted l_i, the segment l_i connecting the articulations P_i and P_{i+1}. The position of the final body corresponds to the wrist position, i.e. with P_3, end of the segment l_2. Since those articulations allows only pure rotations of the segment l_i around the articulation P_i, we can define the set of the reachable positions by the articulation $P_j (j = 2, 3)$ as a sphere, centered on the preceding articulation P_{j-1} and whose ray is $\|l_i\|$[3] (cf. figure 1). Using

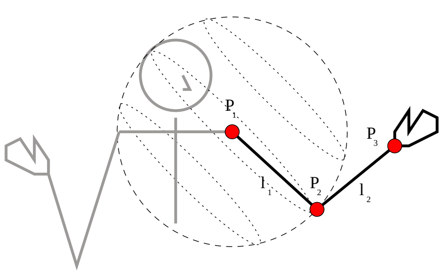


Figure 1: Model of the articular system of the arm. The sphere represents the set of the possible positions for the elbow.

this model, the reachable space for each articulation position becomes a sphere whose parameters are known if we determine the position of the preceding articulation and the length of each segment of the arm, which means that we have to know, from the beginning, the $3D$ position of the shoulder. This can be problematic in an uncontrolled environment. However, when the problem is to obtain qualitative or relational values, or for angular values calculation, a reconstruction with a scale factor can be sufficient. The position of the shoulder could then be fixed as a preliminary. Identically, dimensions of each segment can be fixed arbitrarily as long as the ratio of their respective lengths is respected.

[3] $\|l_i\|$ is the norm of the segment l_i

2.1.3. Algorithm

The method we present exploits a simple geometrical model of the scene and especially of the structure of the arm. We suppose that the coordinates of the points corresponding to the articulations are known, in the image. They can be writen in homogeneous coordinates as :

- for the shoulder : $\widetilde{p_1} = (u_1, v_1, 1)^T$
- for the elbow : $\widetilde{p_2} = (u_2, v_2, 1)^T$
- for the wrist : $\widetilde{p_3} = (u_3, v_3, 1)^T$

After the calibration of the camera, we can compute for each point in the image the associated projection ray, which is the line (passing by the optical center and the point image considered) containing the $3D$ counterpart of this point.

Set of possible configurations for the elbow : knowing P_1 the (possibly arbitrary) position of the shoulder in space, the set of possible positions for the elbow can be defined as the sphere S_1 centered on the shoulder and whose ray is the length $\|l_1\|$ of the arm. The Cartesian equation of such a sphere is :

$$(X_1 - x)^2 + (Y_1 - y)^2 + (Z_1 - z)^2 - \|l_1\|^2 = 0 \quad (1)$$

Equation of the projection ray : $\widetilde{p}_1$ is the position of the shoulder in the image, expressed in homogeneous coordinates. The calibration of the camera gives us the perspective transformation C matrix defining the transformation from a $3D$ frame associated to the camera[4], to the $2D$ image frame [5]. The matrix defining the perspective transformation which forms the image is traditionnaly written as follows :

$$\mathbf{C} = \begin{bmatrix} fk_u & 0 & u_0 \\ 0 & fk_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

Where:

- f is the focal length ;
- k_u, k_v are scale factor, horizontal and vertical, in $pixels/mm$
- (u_0, v_0) is the position of the principal point in the image frame (the projection of the optical center of the camera).

[4] the origin of this frame is in the optical center

[5] the origin of the image frame is in the left higher corner of the image

This matrix let us deduct the position in the image frame of a point $\widetilde{\mathrm{p}}_{\mathrm{i}} = (u_i, v_i, 1)^T$ projection of a point whose coordinates are expressed in the camera frame $P_i = (X_i, Y_i, Z_i)^T$:

$$\begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \begin{bmatrix} fk_u & 0 & u_0 \\ 0 & fk_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} \quad (3)$$

The inverse of this matrix is used to calculate, for each point p_i in the image, the equation of the associated projection ray in space. The projection ray is the line passing through the focal point of the camera and the considered point in the image plane. The original $3D$ point is necessarily located on this line. Here is a parametric equation of the projection ray, where λ is a simple multiplying coefficient :

$$R_i(\lambda) = \lambda \widetilde{\mathbf{p}}_{\mathbf{i}} \quad (4)$$

$\widetilde{\mathbf{p}}_{\mathbf{i}}$ represents the coordinates of the image point in the camera frame :

$$\widetilde{\mathbf{p}}_{\mathbf{i}} = C^{-1}\widetilde{\mathrm{p}}_{\mathrm{i}} \text{ with } C^{-1} = \begin{bmatrix} \frac{1}{fk_u} & 0 & \frac{u_0}{fk_u} \\ 0 & \frac{1}{fk_v} & \frac{v_0}{fk_v} \\ 0 & 0 & 1 \end{bmatrix} \quad (5)$$

$$\text{So that : } \widetilde{\mathbf{p}}_{\mathbf{i}} = \begin{bmatrix} \frac{(u_i - u_0)}{fk_u} \\ \frac{(v_i - v_0)}{fk_v} \\ 1 \end{bmatrix} \quad (6)$$

Therefore the $3D$ position we search is the intersection of the surface of the sphere S_1 defining the set of the possible configurations for the elbow, and the projection ray $r_i(\lambda)$. Calculation of those intersections in the camera frame consists in determining values for λ such as :

$$\begin{aligned} &(X_1 - \lambda\tfrac{(u_i - u_0)}{fk_u})^2 + (Y_1 - \lambda\tfrac{(v_i - v_0)}{fk_v})^2 \\ &+(Z_1 - \lambda)^2 - \|l_1\|^2 = 0 \end{aligned} \quad (7)$$

This is a second degree polynomial $a\lambda^2 + b\lambda + c = 0$ whose coefficients are :

$$\begin{aligned} a &= (\tfrac{(u_i - u_0)}{fk_u})^2 + (\tfrac{(v_i - v_0)}{fk_v})^2 + 1; \\ b &= 2[(\tfrac{(u_i - u_0)}{fk_u})(-X_1) + (\tfrac{(v_i - v_0)}{fk_v})(-Y_1) - Z_1]; \\ c &= X_1^2 + Y_1^2 + Z_1^2 - l_1^2 \end{aligned} \quad (8)$$

Solving this polynomial gives two possible values for λ, possibly a single double one, the positions $\widehat{p}_{2,j} (j = 1, 2)$ possible for the elbow comes now directly since $r(\lambda) = \lambda\widetilde{\mathbf{p}}_{\mathbf{i}}$.

Using the same technique, we are able to calculate the possible positions $\widehat{p}_{3,j} (j = 1..4)$ of the wrist, considering the two spheres whose centers are given by the estimated positions of the elbow and rays by the length of the forearm. We can calculate for each value of the position of the elbow two possible positions for the wrist and thus four possible configurations for the arm.

This algorithm allows us to reduce the set of possible configurations for an arm to four possibilities for a single image. Elbow's positions are symmetric in regard of a plane parallel to the image and containing the shoulder. Calculation of the wrist's position is performed from each possible elbow position so that we obtain four possible positions for the wrist. In the same way as for the elbow, each couple of solutions is symmetric in regard of a plane parallel to the image and containing the corresponding elbow position.

2.2. Extension to image sequences analysis

In the case of image sequences, we calculate a set of $3D$ points candidates for each image. During the sequence those points have to be merged to build trajectories. For each branch of the solution tree (except particular configuration) there are two points to assign to a pair of trajectories. Since it is not possible to know directly which point must be attached to a given trajectory, we introduce a linearity criterion : we calculate the angle α between vectors $\overrightarrow{\mathcal{V}}_{i,j,k}$ and $\overrightarrow{\mathcal{V}}_{i,j,k+1}$, where $\overrightarrow{\mathcal{V}}_{i,j,k}$ is defined by points $\widehat{P}_{i,j,k-1}$ and $\widehat{P}_{i,j,k}$, and $\overrightarrow{\mathcal{V}}_{i,j,k+1}$ by points $\widehat{P}_{i,j,k}$ and $\widehat{P}_{i,j,k+1}$. $\widehat{P}_{i,j,k}$ is the $j^{th} (j = 1, 2)$ estimated space coordinates of the articulation i in the k^{th} image of the sequence. We must therefore calculate the norm of the cross product $\|\overrightarrow{\mathcal{V}}_{i,j,k} \wedge \overrightarrow{\mathcal{V}}_{i,j,k+1}\| = \|\overrightarrow{\mathcal{V}}_{i,j,k}\|\|\overrightarrow{\mathcal{V}}_{i,j,k+1}\| sin(\alpha)$. (figure 2).

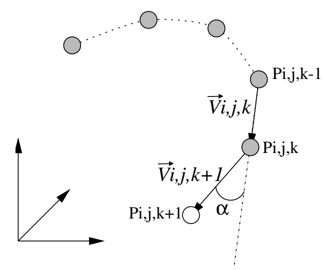


Figure 2: Building trajectories : we first compute the cross product between the last guiding vector of the current trajectory and the new one build by using the (white) candidate point. Linearity criterion consists in merging to the current trajectory the point which minimises the norm of this cross product.

The candidate for which the norm is weakest is affected to the corresponding trajectory. The second point of the branch is then affected to the other one.

Particular configurations : The construction of the trajectories described above can be done correctly in the general case where the algorithm gives two intersections between the projection ray and the sphere. However there are configurations where this assumption is false. Those configurations must be taken into account in the algorithm ; they can also be used as detector for particular movements. There are two categories of particular configurations:

1. The polynomial (8) has only a single solution. It happens when the considered segment (arm or forearm) is included in a plane parallel to the image plane. In this case, the projection ray is tangent to the sphere and there will only be a single "intersection" with the sphere. This point is then added to the both trajectories : it indeed corresponds to a case where the two possible trajectories will cross.

2. The polynomial (8) does not have any solution. In the absence of noise, this case can occur only for the wrist : after having calculated the two possible positions for the elbow, we define the pair of spheres which forms the set of the possible positions of the wrist. There are cases where, based on the "wrong" position of the elbow, the sphere does not have any intersection with the

i	θ_i	d_i	α_i	a_i
1	θ_1	0	$-\pi/2$	0
2	θ_2	0	$\pi/2$	0
3	θ_3	l_1	$-\pi/2$	0
4	θ_4	0	$\pi/2$	0
5	0	l_2	0	0

Table 1: DH parameters describing the human arm system

projection ray. Those configurations directly allows us to cut a complete branch from the solution tree.

2.2.1. Angular values calculation

The parametric model of the human arm is based on the modified Denavit-Hartenberg (DH) parameters description (Denavit J., 1955). This representation provides a systematic method for describing relationships between adjacent links. As long as the frames attached to each articulation are positionned using DH algorithm (Cf. 2.2.1.. The model consists in a $4x4$ homogeneous transformation matrix corresponding to the transformation from link 1 to link 3, which describes, in fact, the arm system. This matrix is parametrized with angular values of each joint and link lengths. This matrix constitute the direct geometrical model. Whereas the inverse geometrical model provides the joint angular values in function of the joint Cartesian coordinates.

Modified parameters of Denavit-Hartenberg : The DH method is systematic as long as the axis system R_i attached to each joint (figure 3) is defined using the following rules :

1. O_{i-1} is the perpendicular common to link L_{i-1} and L_i axes located on link L_{i-1} ;
2. axis x_{i-1} is the unit vector of the common perpendicular oriented from link L_{i-1} to link L_{i-1} ;
3. z_i is the unit vector of link L_i ;
4. axis y_i is set so that : $y_i = z_i \wedge x_i$
5. relationships between frame R_i and R_{i-1} are defined by the following parameters :
   - α_i is the offset angle from axis z_{i-1} to z_i around x_{i-1} ;
   - d_i : the distance from the origin of the $(i-1)^{th}$ coordinate frame to the intersection of the z_{i-1} axis with the x_i ;
   - θ_i : the joint angle from x_{i-1} to x_i turning around z_i ;
   - a_i : the offset distance from the intersection of the z_{i-1} axis with the x_i axis.

With the joint frame O and 1 jointed, the arm model is given by the D-H parameters is shown in table 1.

DH parameters are used to write an homogeneous transformation matrix for each joint. The generic form of the matrix for a revolving joint is :

$$^{i-1}T_i = \begin{bmatrix} \cos\theta_i & -\sin\theta_i & 0 & a_i \\ \cos\alpha_i \sin\theta_i & \cos\alpha_i\cos\theta_i & -\sin\alpha_i & -d_i\sin\alpha_i \\ \sin\alpha_i\sin\theta_i & \sin\alpha_i & \cos\theta_i & d_i\cos\alpha_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (9)$$

Where $\theta_i, \alpha_i, d_i, a_i$ are the DH parameters.

Direct geometrical model : The direct geometrical model gives the transformation from Cartesian coordinate space to angular values of the each joint. The $4x4$ matrix $^0T_5$ specifies the homogeneous transformation from frame 0 (the shoulder) to frame 5 (the wrist) (figure 3). This matrix is built by multiplying the successive homogeneous transformation matrices $^{i-1}T_i, i = 0, .., 5$.

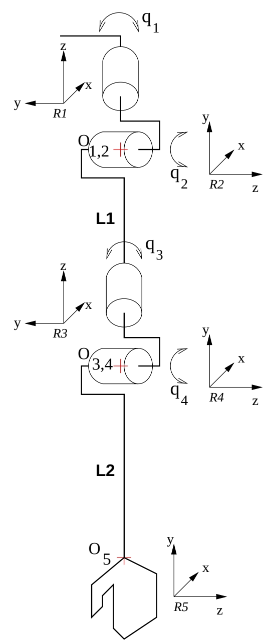


Figure 3: Arm model showing the frames used in direct geometric model calculation

This model is parametrized by θ_i, the joints angular values, and allows cartesian coordinates calculation. For angular coordinates calculation we need to inverse this model.

Inverse geometrical model : The inverse geometrical model is parametrized by the Cartesian coordinates of the wrist and returns the angular value θ_i for each joint. The first way to calculate this model would be to calculate the inverse of $^0T_5$, but in regard of the complexity of the calculation, splitting up the kinematic chain will be a far better solution. We calculate angular values for each joint separately by defining the inverse transformation $^3T_0$ that gives us the shoulder's joints angular values from elbows's Cartesian coordinates (expressed in frame R_0) and $^5T_3$ which gives us elbow's angular values from wrist's Cartesian coordinates expressed in R'_2 frame. R'_2 is a virtual frame oriented as R_2 and centered on the elbow.

Considering only the shoulder, we can write DH parameters (table 2) for the shoulder-elbow system, and define the homogeneous transformation matrix $^0T_3$ transformation matrix by multiplying the elementary transformation matrices (9)which specifies the transformation from frame

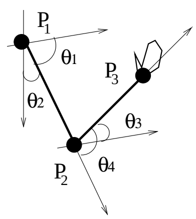


Figure 4: The inverse geometric model of the arm gives the joint angular values θ_1, θ_2 (shoulder joint) , θ_3, θ_4 (elbow joint) knowing shoulder, elbow and wrist Cartesian coordinates P_1, P_2, P_3

i	θ_i	d_i	α_i	a_i
1	θ_1	0	$-\pi/2$	0
2	θ_2	0	$\pi/2$	0
3	0	$-l_1$	0	0

Table 2: DH parameters for shoulder-elbow system

R_{i-1} to frame $R_i (i = 1, 2, 3)$:

$$^0T_3 = \begin{bmatrix} \cos\theta_{12} & -\sin\theta_1 & \cos\theta_{12} & -l_1\cos\theta_1\sin\theta_2 \\ \sin\theta_1\cos\theta_2 & \cos\theta_1 & \sin\theta_{12} & -l_1\sin\theta_{12} \\ -\sin\theta_2 & 0 & \cos\theta_2 & l_1\cos\theta_2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Where $\cos\theta_{12}$ stands for $\cos\theta_1 * \cos\theta_2$.
The fourth column of $^0T_3$ represents the direct geometric model, so the inverse geometric model is :

$$\begin{cases} \theta_1 = \arctan(y_2/x_2) \\ \theta_2 = \arccos(z_2/l_1) \end{cases} \quad (10)$$

Doing the same calculation for the wrist brings (wrist's cartesian coordinates have to be expressed in elbow-centered frame R'_2) :

$$\begin{cases} \theta_3 = \arccos(z_3/l_2) \\ \theta_4 = \arctan(y_3/x_3) \end{cases} \quad (11)$$

As the arm model is redundant, direct inversion using the analytical solution will lead to unexpected reversal in angular values. To avoid it, we use a numerical resolution method to compute the first two joint values (θ_1, θ_2). This method can be initialized with previously computed values so that the new ones stay as close as possible to them which leads to smooth trajectories. Solution is computed iteratively using the pseudo-inverse of the arm system Jacobian (Klein C.A., 1983). Only the last two values are analytically calculated. This approach allows us to obtain a set of angular values corresponding to the given $3D$ joints position, even when the arm has a singular configuration.

3. Articulation-based motion representation

Articulation-based motion representation could be used to distinguish, among geometrical solutions, the good one, so that the first point to study is the variation of joint values for each solution. The second one concerns the possibility to use those representations to differentiate gestures based on the way they are made. Preliminary experiences have been made using a video corpora of elementary gestures.

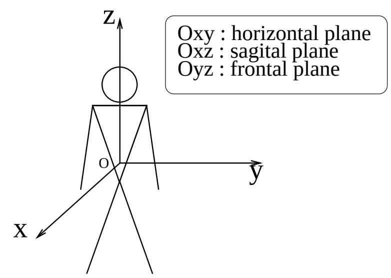


Figure 5: Representation of speaker-centered frame (ScF) showing the planes where most of the elementary gestures are realized.

The results (fig. 6) concern two circular motions of the left hand done in a plane parallel to Oyz plane of ScF, the first one with the arm in extension (gesture **A**) and the other one with the elbow bended (gesture **B**). The third gesture presented is a circular one made with the elbow bended in a plane parallel to Oxz plane of ScF (fig. 5) (gesture **C**). So that gestures **A** and **B** have quite similar aspect from the viewer's point of view and that gesture **B** and **C** are performed by moving articulations in a similar manner. Joints values are computed on each solution provided by the geometric reconstruction algorithm.

Figure (6) presents angular values evolution for each joint of the arm model and for three different gestures. Those values are presented in polar coordinates and ρ parameter stands for time (which means that gesture duration has been normalized). Different curves correspond to angular values computation for each geometrical solution. If we except noise on angular values implied by geometrical reconstruction, different angular trajectories for a same angle can be either confused (fig.6, θ_1 and θ_4 variations for gesture **A**) or symmetric (fig.6, θ_1 and θ_2 variations for gesture **B** and **C**). So that for each solution, changes in angular value variation occur at the same time.

One can remark too, that gesture **B** and **C** have closer signatures than gesture **A** and **B** in the sense that θ_1,θ_2 and θ_3 variations have the same kind of symmetry for those gestures : θ_1 and θ_2 are symmetric in regard of on a horizontal axis and θ_3 values present symmetries in regard of a vertical one. And that angular values for each articulation take values in the same part of the angular space.

4. Conclusion

Articulation-based motion representations are used to improve results computed by a single-camera geometrical algorithm which estimates possible poses of a human arm, being given a single image. This algorithm, provides us with a set of four possible motions for the arm in space. We made the assumption that using such a representation of gesture could allow us to use any of those solutions for gesture analysis. Primary experimentations on simple gestures brought out relationships such as symmetries or confusion between angular values for the different solutions, which is due to symmetries between the different solutions. On the other hand, recent linguistic issues made the assumption that using a proprioceptive representation of gesture is more suitable for Sign language analysis than a description based on elementary gestures described from an ob-

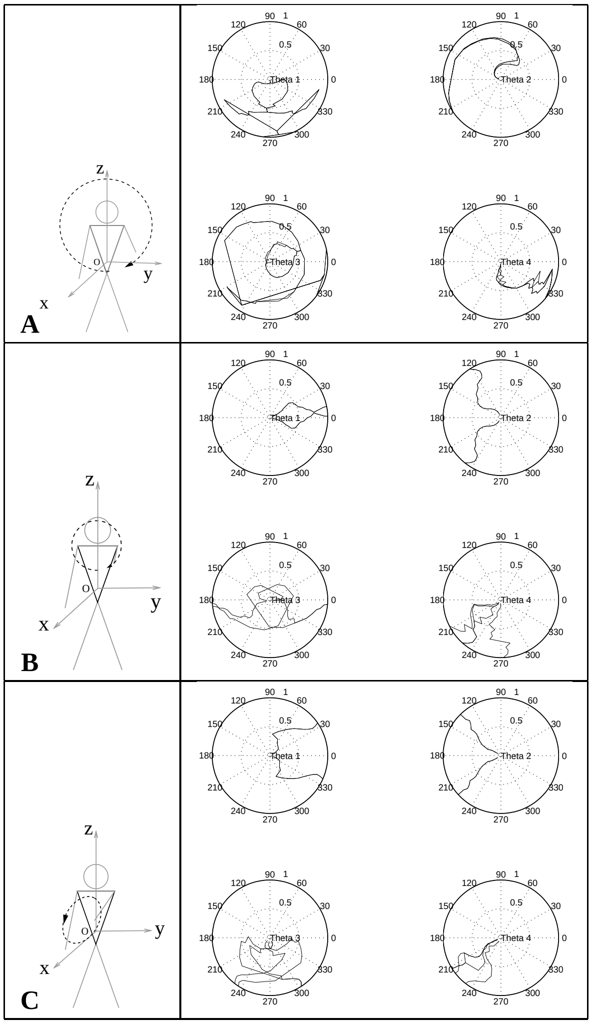


Figure 6: *On the left :* gesture **A,B,C** representation in ScF. Gesture **A** and **B** have similar visual aspect from the viewer's point of view, while gesture **B** and **C** are performed with similar articulation motion. *On the right :* joint values computed on each gesture and for each solution provided by geometrical reconstruction. Each solution is displayed as different curve. Each graph presents the evolution of the angular value for a given angle, from left to right, from top to bottom : $\theta_1, \theta_2, \theta_3, \theta_4$. Angle values are displayed in polar coordinates and ρ parameter stands for the time so that a constant angle value for an angle would be displayed as straight line starting at the center of polar frame.

servator point of view. Our algorithm make it possible to build such a articulation-based motion representation from single-camera data. Considering gestures performed in a similar manner with different orientations and comparing the results to gestures performed in a different manner but similar form observers point of view, we could observe that using our method will lead to a different gesture classification than the ones based on visual aspect in image or tridimentionnal representations. Further researchs have to be perform to bring out useful criterions to analyze real Sign language gestures from this point of view, but primary results are encouraging.

5. References

Athitsos V., Sclaroff S., 2004. Database indexing methods for 3d hand pose estimation. In *Gesture Workshop*, volume 2915 of *Lecture Notes in Computer Science*. Genova, Italy: Springer.

Boutet, D., 2001. Une approche morphogénétique du sens dans la gestuelle conversationnelle. In C. Calvé et al. (ed.), *Oralité et Gestualité*. Paris: l'Harmattan.

Braffort, A., 1996. Argo : An architecture for sign language recognition and interpretation. In P. Harling and al. (eds.), *Progress in Gestual Interaction*. Springer.

Denavit J., Hartenberg R.S., 1955. A kinematic notation for lower-pair mechanisms based on matrices. *Journal of Applied Mechanics*:215–221.

Gurdjos P., Payrissat R., 2002. Calibrage plan d'une caméra en mouvement à focale variable. In *RFIA 2002*. AFRIF-AFIA.

Heikkilä, J., 2000. Geometric camera calibration using circular control points. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1066–1077.

Horain P., Bomb M., 2002. 3d model based gesture acquisition using a single camera. In *Proceedings of the IEEE Workshop on Applications of Computer Vision (WACV 2002)*. Orlando.

Klein C.A., Huang C.H., 1983. Review of pseudoinverse control for use with kinematically redundant manipulators. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-13(3):245–250.

Lenseigne B., Gianni F., Dalle P., 2004. Estimation monovue de la posture du bras, méthode et évaluation. In *RFIA 2004*. Toulouse: AFRIF-AFIA.

Prillwitz, S. and al., 1989. *HamNoSys. Version 2.0; Hamburg Notation System for Sign Languages. An introductory guide*. Hamburg : Signum.

Somers M.G., Whyte R.N., 2003. Hand posture matching for irish sign language interpretation. In *ACM Proceedings of the 1st international symposium on Information and communication technologies*. Dublin Ireland: Trinity College Dublin.

Starner T., Pentland A., Weaver J., 1998. Real-time american sign language recognition using desk and wearable computer based video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(12):1371–1375.

Tanibata N., Shirai Y., Shimada N., 2002. Extraction of hand features for recognition of sign language words. In *15th International Conference on Vision Interface*. Calgary, Canada.

Vogler C., Metaxas D., 1998. Asl recognition based on a coupling between hmms and 3d motion analysis. In *Proceedings of the International Conference on Computer Vision*. Mumbai, India.

Vogler C., Metaxas D., 1999. Parallel hidden models for american sign language recognition. In *Proceedings of the International Conference on Computer Vision*.

Wren C., Pentland A., 1999. Understanding purposeful human motion. In *Proc. IEEE International Workshop on Modelling People (MPEOPLE)*.

Phonetic Model for Automatic Recognition of Hand Gestures

Jose L. Hernandez-Rebollar
The George Washington University, Electrical and Computer Engineering
725 23rd St. NW. Lab 302, Washington DC 20052
jreboll@gwu.edu

Abstract

This paper discusses a phonetic model of hand gestures that leads to automatic recognition of isolated gestures of the American Sign Language by means of an electronic instrument. The instrumented part of the system combines an AcceleGlove and a two-link arm skeleton. The model brakes down hand gestures into unique sequences of phonemes called Poses and Movements. Recognition system was trained and tested on volunteers with different hand sizes and signing skills. The overall recognition rate reached 95% on a lexicon of 176 one-handed signs. The phonetic model combined with the recognition algorithm allows recognition of new signs without retraining.

1. Introduction

The development of automatic means to study sign languages is destined to have enormous impact on economy, society and science. Costello [1999] estimates that American Sign Language (ASL) is the fourth most used language in the United States with 13 million people, including members of both the hearing and deaf community. Some 300,000 to 500,000 of them are ASL native-speakers, which means that their full integration to society depends on their ability to overcome the language barrier by using all means at their disposal. William Stokoe [1995] was probably the first linguist to involve engineers, not only educators, in solving the challenge of better communication, he wrote: "Looking back, it appears that linguistics was made possible by the invention of writing. Looking ahead, it appears that a science of language and communication, both optic (gestures) and acoustic (speech), will be enabled, in all probability, not by refinements in notational systems, but by increasing sophistication in techniques of recording, analyzing, and manipulating visible and auditory events electronically."

It is ironic that even though humans learned how to communicate through gestures before learning how to speak, methodologies for analyzing speech and spoken languages are far better understood than the methodologies for analyzing and, in consequence, recognizing gestures and sign languages.

Engineers found a way to capture speech in 1915 with the invention of the carbon microphone. This transducer produces an electrical signal corresponding to change in air pressure produced by sound waves, which contains all the information required to record and reproduce speech through a speaker. Sign language, in turn, combines hand movements, hand shapes, body posture, eye gaze, and facial expression that are not easy to capture by using only one type of sensor. Approaches that use arrays of video cameras to capture signing struggle to find an adequate way of reproducing tri-dimensional images. The high resolution needed to capture hand shape and eye gaze results in a reduced field of view unable to fit hand movement or body posture, and a high bandwidth connection (processor) is required to transmit (analyze) the data stream and reproduce the video at acceptable speed.

An alternative is the combination of angular sensors of different types mounted directly on the signer's joints of interest. Although bulkier, cumbersome and more obtrusive, these instrumented approaches have been more successful in capturing hand postures [Grimes1983, Kramer1998] than the approaches based on video alone [Uras1994].

In this work the combination of a phonetic model of hand gestures and a novel instrumentation to capture and recognize the hand gestures in American Sign Language, is discussed. Non-manual components such as facial expression, eye gaze and body posture are not considered here.

2. Review of previous approaches.

The first and most important step in the recognition process is to extract, from a given gesture, all the necessary features that allow the recognition system to classify it as member of one and only one class. Two things are needed to achieve that step: a model that describes gestures in terms of necessary and sufficient features, and a capturing system suitable to detect such features. It is imperative for the resulting set of features (*pattern*) to be different for each gesture, and it is desirable for the resulting pattern to have a constant number of features (*fix dimensionality*) and as few as possible (*reduced dimensionality*).

The model proposed in this work is based on the assumption that any hand gesture can be analyzed as a sequence of simultaneous events, and each sequence is unique per gesture. Those events are referred in this work as phonemes. As straightforward as this scheme could sound, it could be cause of debate among many signers and teachers who conceive signs as indivisible entities. The following is a review of different phonemes and structures that have been proposed to model hand gestures.

2.1. Phonetic structure

By using traditional methods of linguistics to isolate segments of ASL, Stokoe found that signs could be broken down into three fundamental constituent parts: the hand shape (*dez*), hand location with respect to the body (*tab*), and the movement of the hand with respect to the body (*sig*), so these phonemes happen simultaneously. Lidell [1989] proposed a model of movements and holds, Sandler [1986] proposed movements and locations, and Perlmutter

[1988] proposed movements and positions, all of them happening sequentially.

Some automatic systems have followed models similar to Stokoe [Bauer, 2000; Vamplew, 1996] and Lidell [Vogler, 1999]. By using Stokoe's model, patterns are of reduced and fix dimensionality but similar for gestures that are only different in their final posture (such as GOOD and BAD). Patterns that result from Liddell's model eliminate this problem by considering the initial, final, and intermediate states of the hand and the movements that happen in between. Still, the model produces ambiguous patterns with variable dimensionality. As an example, when signing FATHER, tapping the thumb of a 'five' hand shape against the forehead, the sequence can be described as a Movement followed by a Hold followed by a Movement and finished by a Hold (MHMH) or as a HMHMH if the hand is considered to start from a static position, or as a simple Hold, as many signers do not make long movements when tapping. Closely linked to these models are the recognition methods suitable to recognize the resulting patterns. Hidden Markov Models (HMM) and Neural Networks (NN) have been used to recognize complete sentences [Starner, 1998], isolated words [Waldron, 1995], or phonemes [Vamplew, 1996], but none of those approaches has been able to integrate hand gesture and finger spelling in one recognition system.

2.2. The Pose-Movement model

Under the sequential models previously explained, ASL resembles the linear structure of spoken languages: phonemes make up words, and words in turn make up sentences. Phonemes in these models are, in some degree, the three simultaneous components of Stokoe, so the execution of ASL gestures can be seen as a sequential combination of simultaneous phonemes. Specifically, two types of phonemes: one static and one dynamic.

Definition 1: A ***pose*** is a static phoneme composed of three simultaneous and inseparable components represented by vector **P** = [hand shape, palm orientation, hand location]. The static phoneme occurs at the beginning and at the end of a gesture.

Definition 2: A ***posture*** is a vector of features **Ps** = [hand shape, palm orientation]. Twenty-four out of the 26 letters of the ASL alphabet are postures that keep their meaning regardless of location. Letters J and Z are not considered postures because they have movement.

Definition 3: ***Movement*** is a dynamic phoneme composed by the shape and direction of the trajectory described by hands when traveling between successive poses. **M**=[direction, trajectory].

Definition 4: A ***manual gesture*** is a sequence of poses and movements, P-M-P.

Definition 5: **L**, the set of purely manual gestures that convey meaning in ASL is called the ***lexicon***.

Definition 6: A manual gesture **s** is called a ***sign*** if s belongs to L.

Definition 7: ***Signing space*** refers to the physical location where signs take place. This space is located in front of the signer and is limited by a cube bounding the head, back, shoulders and waist.

By following definitions 1 to 7, icons, letters, initialized, and non-initialized signs, are modeled by PMP of fixed dimensionality, while compound, pantomimic, classifiers, and lexicalized finger spelled words, are modeled as sequences of variable length. These patterns are listed in Table 1.

Sign	Model
Two handed icons	PMP, PMP one sequence per hand
Finger spelled words	PMP per letter
Lexicalized finger spelled, *compound signs, **pantomimic	sequence of 2n-1 phonemes n= number of letters n=number or signs* n=number of pauses**

Table 1. Signs and their respective sequences of phonemes

As a proof of concept, a Lexicon of one-handed signs from two dictionaries [Costelo,1999; IDRT, 2001] with patterns of the form PMP were targeted for recognition. Since any sign is merely a new combination of the same phonemes, the recognition system is composed by small subsystems that capture a finite number of phonemes complemented by a search engine, which compares captured sequences against stored sequences.

3. Instrumentation

The instrument designed to capture all the phonemes found in the resulting sequences (53 postures, including six orientations; twelve movements and eleven locations) comprises an Acceleglove [Hernandez, 2002] to capture hand postures, and a two-link skeleton attached to the arm to capture hand location (with respect to the shoulder) and hand movement. Data is sent serially to a laptop Tthinkpad running windows 98 on a Pentium III. The sign recognizer is based on a search algorithm.

3.1 Training and testing

Posture, location and movement were recognized independently; trained and tested with help of 17 volunteers of different skill levels, from novice to native signer. That selection allowed covering a range of accents and deviations with respect to the citation form. The search algorithm was tested with 30 one-hand gestures first, and 176 later to test scalability. The complete list of signs is found in [Website].

3.2. Postures

The posture module starts recognizing any of six palm orientations: vertical, horizontal, vertical up-side down, horizontal tilted, horizontal palm up, and horizontal tilted counter clockwise.

Afterwards, the posture recognizer progressively discriminates postures by the position of fingers. Decision trees are generated as follows [Hernandez, 2002b].

-**For** all trees, **start** decision nodes evaluating the position of the pinky finger and base the subsequent node's decision on the next finger (ring, middle, index, thumb).

-**If** postures are not discriminated by finger flexion, **then** continue with finger abduction.

-**If** postures are not different by individual finger flexions or abductions, **then** base classification on the overall finger flexion and overall finger roll.

To train the orientation nodes, all seventeen signers were asked to hold the initial pose of FATHER, NICE, PROUD, PLEASE, THING and ASIDE. In average, the orientation module accurately recognized 94.8% of the samples. The worst recognition rate corresponded to horizontal postures where the threshold is blurred by the deviations introduced by signers' accents, since they were asked to hold their poses, not to hold their hand in a certain position.

3.2.1. Aliases

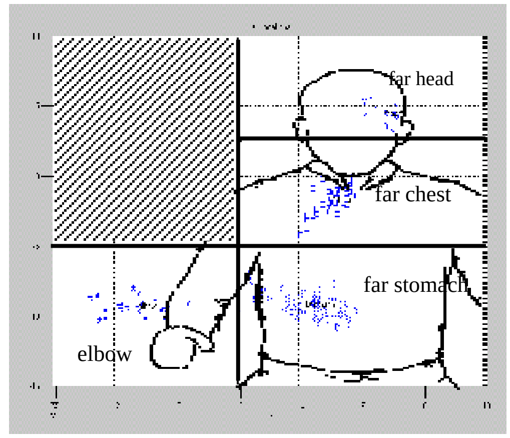
Since accelerometers do not detect angular positions around the gravity vector, 10 postures were impossible to discriminate based on finger bending or abduction around the gravity vector. These postures are called *aliases*. This aliasing reduced the number of recognizable postures from 53 to 43. The highest accuracy (100%) corresponded to the vertical palm with knuckles pointing down used to sign PROUD, the worst accuracy rate corresponded to postures C and E, with 68%, for a recognition average of 84%.

3.3. Locations

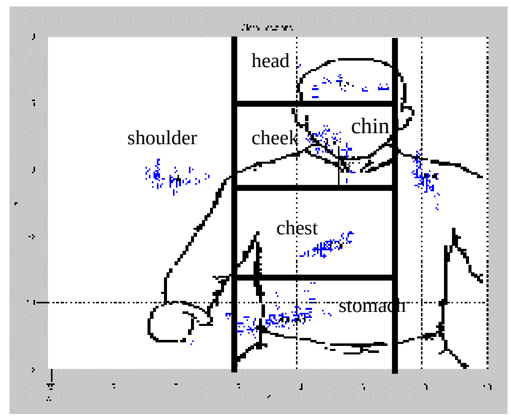
By looking at the initial and final position of the hand during the execution of each sign in the lexicon, eleven regions in the signing space were identified: head, cheek, chin, right shoulder, chest, left shoulder, stomach, elbow, far head, far chest and far stomach. To train the recognizer, four signers were asked to locate their hand at the initial poses of several signs that start or finish at those regions: FATHER, KNOW, TOMORROW, WINE, THANK YOU, NOTHING, WHERE, TOILET, PLEASE, SORRY, KING, QUEEN, COFFEE, PROUD, DRINK, GOD, YOU, FRENCH FRIES and THING. Volunteers were chosen based on their heights so they cover the full range of height among the group of volunteers.

Figure 1 shows the initial and final locations captured with the two-link skeleton as executed by the middle height signer (1.70 mts). Figure 1a corresponds to locations close to the body and Figure 1b corresponds to locations away from the body. A human silhouette is superimposed on the plane to show locations related to signer's body. The plane *y-z* is parallel to the signer's chest, with positive values of *y* running from the right shoulder to the left shoulder and positive values of *z* above the right shoulder.

Similar to orientations and postures, locations are solved using a decision tree, thresholds on *y* and *z* boundaries are set at least 4σ around the mean, and 3σ on *x* due limitations imposed by the instrumentation.



(a)



(b)

Figure 1. a) Far locations. b) Close locations.

The overall accuracy rate was 98.1% : head 98%, cheek 95.5%, chin 97.5%, shoulder 96.5%, chest 99.5%, left shoulder 98.5%, far chest 99.5%, elbow 94.5 %, stomach, far head and far stomach 100%. The skeleton system does not need an external reference source, and it is immune to ambient noise; that makes it a better choice for a portable instrument that infrared and magnetic trackers.

3.4. Movements

Movements of the one-handed signs considered in this work are described by means of two movement primitives: curviness [Bevilaqua2001] and direction. Both metrics are orientation and scale independent. As with the case of hand postures and locations, the exact movement varies from signer to signer and from trial to trial. Six directions (up, down, right, left, towards, and away) and two levels of curviness (straight and circular) were identified in the Lexicon that gave a total of twelve different movements. Same four signers were asked to perform the six basic movements along the main axes and the two curves ten times each. Directions ‘left’ and ‘right’ were classified with less than 100% (77% and 75%) reducing overall accuracy to 92%. A curviness greater than 4 discriminated circles from straight lines with 100% accuracy, but only signs with straight movements were implemented in the recognition algorithm.

4. Search Engine.

A variation of template matching called *conditional template matching* was used to classify complete signs. Conditional template matching compares the incoming vector of phonemes (captured with the instrument) against a pre-stored file of patterns, component by component, and stops the comparison when a condition is met:

-**For** all patterns in the lexicon, extract a list of signs matching the initial **posture** captured by the Acceleglove. This is the first list of candidate signs.

-**For** all patterns in the list of candidates, select the signs matching the **initial location** captured by the two-link skeleton. This is the new list of candidate signs.

Repeat the matching and creation of new lists of candidates by using movement, final posture and final location.

Stop when all components have been used **OR** when there is only one sign on the list after matching the initial location. That sign on the list is called 'the most likely'.

The search algorithm can be seen as a decision tree with a variable number of nodes. The expected probability of finding a given sign is inversely proportional to the depth of the tree. In other words, it is more likely to recognize a sign if it is the only one in the lexicon performed with certain initial pose (such as PROUD), and it is less likely to recognize two signs when only the final pose makes them different (such as GOOD and BAD).

4.1. Evaluation.

An initial evaluation used only 30 signs taken from Starner (1998), Vogler (1999), and Waldron (1995): BEAUTIFUL, BLACK, BROWN, DINNER, DON'T LIKE, FATHER, FOOD, GOOD, HE, HUNGRY, I, LIE, LIKE, LOOK, MAN, MOTHER, PILL, RED, SEE, SORRY, STUPID, TAKE, TELEPHONE, THANK YOU, THEY, WATER, WE, WOMAN, YELLOW, and YOU. The PMP sequences reflect the citation forms as found in Costello [1999] and in the Ultimate ASL Dictionary [IDRT2001]. The overall recognition rate was 98% since almost all of them have different initial poses.

4.2. Scalability

Since any new sign is a combination of the same phonemes, the lexicon can be expanded without retraining the search algorithm. When tested on 176 one handed signs performed by one signer the overall recognition rate reached 95%.

5. Conclusions and Future Work

The model, instrumentation and recognition algorithm explained in this work represent a framework for a more complex system where a larger lexicon can be recognized by extending the patterns to include non-manual gestures when the required instrumentation to detect them becomes available.

Work in the immediate future will incorporate a second PMP sequence for the non-dominant hand, and migrate the recognition program to a wearable computer for a truly portable electronic translator. The long-term objective shall include a grammar correction module to rearrange the sequence of translated glosses and correct for tenses, gender, and number as needed by the spoken language.

6. References

- Bauer, B., Hienz, H., and Kraiss, K., 2000. Video-Based Continuous Sign Language Recognition Using Statistical Methods. IEEE 2000, pp 463-466.
- Bevilacqua F., Naugle L., and Valverde I., 2001. Virtual Dance and Music Environment Using Motion Capture. Proc. of the IEEE-Multimedia Technology and Applications Conference, Irvine, CA.
- Costelo, Eleine 1999. Random House Webster"s Concise American Sign Language Dictionary. Random House Inc. NY.
- Fels, Sidney S., and Hinton, Geoffrey E., 1993. Glove Talk -A Neural-Network Interface Between a Data-Glove and a Speech Synthesizer. IEEE Transactions on Neural Networks, vol. 4, No. 1. January.
- Grimes, G. 1983. US Patent 4,414,537. November.
- Hernandez Jose L., Kyriakopoulos, N., Lindeman, R.. The AcceleGlove a Hole-Hand Input Device for Virtual Reality. ACM SIGGRAPH Conference Abstracts and Applications 2002. pp 259.
- Hernandez, Jose L., Kyriakopoulos, N., Lindeman R. A Multi-Class Pattern Recognition of Practical Finger Spelling Translation, IEEE International Conference on Multimodal Interfaces ICMI'02. October 2002, pp 185-190.
- IDRT 2001. The Ultimate American Sign Language Dictionary. The Institute for Disabilities Research and Training Inc. Copyright 2001.
- Kramer, J., and Leifer, L., 1988. The Talking Glove: An Expressive and Receptive Verbal Communication Aid for the Deaf, Deaf-Blind, and Nonvocal, SIGCAPH 39, pp.12-15 (spring 1988).
- Lidell S., and Johnson, R., 1989. American Sign Language: The phonological base. Sign Language Studies, 64: 195-277.
- Perlmutter, D., 1988. A moisac theory of American Sign Language syllable structure, paper presented at Second Conference on Theoretical Issues in Sign Language Research. Gallaudet University, Washington, DC.
- Starner, T., Weaver, J., and Pentland, A., 1998. A Wearable Computer Based American Sign Language Recognizer, MIT Media Lab. Technical Report 425.
- Stokoe, William C., Armstrong, David F., Wilcox, Sherman E. Gesture and the Nature of Language. Cambridge University Press, 1995
- Uras, C., and Verri, A. 1994, On the Recognition of The Alphabet of the Sign Language through Size Functions. Dipartimento di Fisica, Universitá di Genova. Proceedings of the 12th IAPR Int. Conf. On Pattern Recognition. Conference B: Computer Vision and Image Processing Vol 2,1994. pp 334-338.
- Vamplew, P., 1996, Recognition of Sign Language Using Neural Networks, Ph.D. Thesis, Department of Computer Science, University of Tasmania.
- Vogler, C. and Metaxas, D., 1999, Toward Scalability in ASL Recognition: Breaking Down Signs into Phonemes. Gesture Workshop"99, Gif-sur-Yvette, France, March 17-19.
- Waldron Majula B. 1995, Isolated ASL Sign Recognition System for Deaf Persons, IEEE Trans. On Rehabilitation Engineering vol 3 No. 3 September.
- Sandler, W. 1986, The Spreading hand autosegment of American Sign Language. Sign Language Studies 50: 1-28.
- website: http://home.gwu.edu/~jreboll/signlist.txt

Development of a new „SignWriter“ Program

Daniel Thomas Ulrich Noelpp
d.noelpp@gmx.ch
Moos bei Köniz, Berne, Switzerland
http://www.signwriter.org/

Abstract

The „Sutton SignWriting“ system is a practical writing system for deaf sign languages. The symbols describe shape, location and movement of hands as well facial expressions and other signing information. „SignWriter Java 1.5/Swing“ is being developed as the successor to „SignWriter DOS“, a program for typing and editing „SignWriting“ texts, used by school children, teachers, linguists and Deaf people. The new Java version 1.5 „Tiger“ is used in development and Swing as the graphical user interface.

1. The new program

A „SignWriter Java 1.5/Swing“ program as the successor to „SignWriter DOS“ programmed by Richard Gleaves is being developed in the new Java 1.5 („Tiger“) version using the Swing graphical user interface library. The existing „SignWriter DOS“ program is a simple, yet powerful program for typing and editing „SignWriting“ texts. As many school children, teachers and linguists are already using this program for their everyday work, it is important that the typing conventions are not changed very much. Support for the SGN files („SignWriter DOS“ file format for „SignWriting“ texts) is important as well. In a summary, former users shouldn't need to change their way of working with „SignWriting“ or not very much.

There are some new features, however: A friendlier user interface (thanks to Swing of Java 1.5) is implemented, which is also easier for new users to understand. And because there are different „alphabets“ in use, a multi-alphabet capability seems to be important, too. The old symbols of „SignWriter DOS“ are retrofitted into the framework of the multi-alphabet capability, or expressed in a simpler way: „SignWriter Java 1.5/Swing“ understands the old „alphabet“, but can work with and convert to the new ones. And another important thing is the support for SWML files (an XML file format to store „SignWriting“ texts, developed by Antônio Carlos da Rocha Costa).

It is hoped that the new „SignWriter“ program is accepted by the SignWriting community as the successor to „SignWriter DOS“. Public release is planned for autumn, 2004.

2. About „SignWriting“ and the old program

„Sutton SignWriting“, developed by Valerie Sutton, is a practical writing system which can be used for all the sign languages of the world. The symbols of „SignWriting“ describe the shape, location and movements of the hands, as well as the facial expressions which a signer makes and other signing information. This writing system gives Deaf people the possibility of writing to each other, making notes and reading text written in their native language.

In the eighties, Richard Gleaves developed the first „SignWriter“ program, which made it possible to type „SignWriting“ on the computer. The latest version 4.4 is now eight years old. It is an excellent software from the early days of personal computers, but it has become somewhat outdated. The computer resources at that time were limited and the operating systems were very different from those of today. The user interface no longer meets the expectations which today's users have. One of the biggest drawbacks to this earlier version is that it only runs under a pure DOS system. Modern Mac OS, Windows NT, 2000 and XP all require a DOS virtual machine to start „SignWriter DOS“. There are other shortcomings: Low resolution of the symbols which leads to visible pixelization (zigzag effect on round curves or oblique lines) and inverted display (white on black). These are all reasons why a successor to the SignWriter DOS is urgently needed by the SignWriting community.

3. Demonstration and Discussion

The program is being redeveloped from scratch using the new version 1.5 of Java and with the Swing graphical user interface library. Development is open source. „SignWriter“ is layered onto an alphabet package called „signwriter.alphabet“ which knows about the symbols and is modeled after Sutton's „SymbolBank“. It is hoped that especially the „alphabet“ package can be reused in other projects outside „SignWriter Java“.

The diagram shows some Java interfaces and classes which make up the programmer's interface to the alphabet. This interface is multi-alphabet capable. The programmer loads an **`Alphabet`** object using the **`Factory.loadAlphabet()`** method. From the alphabet one can manage the symbols and base symbols. The package is immutable: once loaded it is impossible to destroy the alphabet by mistake. For symbols within a sign there's another class called **`SignSymbol`** outside the package (not shown in the diagram). There are many more technical details interesting for developers. But because the audience of the demonstration are end users as well, we will stop here.

Please see fig. 1 at the bottom of the paper for an UML class diagram.

The new features of Java 1.5 are used in the program. They are genericity (especially useful for collections of objects like the symbols of a sign, `Sign.getParts()` returns a list of sign parts with the type `List<Part>` for example); the enhanced for loop for an easier iteration through collections and many others. Important for end-users additionally is the improved look-and-feel of Swing which gives Java applications a more modern and friendlier appearance than before.

The demonstration is an opportunity to show and discuss design decisions and diagrams, screenshots and last-minute experiences and to play with the latest development version of the unfinished software. Developers can ask questions about inner workings. End users about the features and the look-and-feel. It is a big opportunity for the team as well! We need the feedback. Without feedback we don't know whether we do the right thing. You have an impact on the development.

Be warned, however. The software is unfinished and not even in alpha stage. Things might not work at all.

4. About Daniel Noelpp

Born Deaf in Switzerland 1970, he attended a residential school for Deaf children near Berne. Later, he was „mainstreamed“ into a school with hearing children. He received his college diploma in 1989. After several years studying at the University of Berne, he worked as a Software Engineer for the same University as well as for several companies in Switzerland. In 2000, he worked for six months as a Software Consultant in Pune, India. At the present time, he is attending HTI (University of Applied Sciences) in Berne and developing „SignWriter Java 1.5/Swing“ at home.

5. The team members

The software is not developed by Daniel Noelpp alone. The other members of the team are Günel Hayirli (HTI student, hearing) and Matthias Noelpp (Electrical engineer, hard of hearing).

6. Donations

We thank Ingvild Roald, for the generous financial support! The project team is working hard with rather limited resources. If you are willing to give a donation to the development, it is appreciated very much. It is planned to put a list of supporters and donators in the About menu of the SignWriter prominently. Would you like to be included in this list? Please contact Daniel Noelpp.

7. References

Sutton, V. (2004) The International Movement-Writing Alphabet – The IMWA Sign-Symbol-Sequence for All Sign Language and Gesture. La Jolla: Deaf Action Committee for SignWriting.

Sutton, V. (2002) Sutton's Sign-Symbol-Sequence 2002. La Jolla: Deaf Action Committee for SignWriting.

Sutton, V. (2002) Sutton's SymbolBank: Sign-Symbol-Sequence 1999 compared to Sign-Symbol-Sequence 1995. La Jolla: Deaf Action Committee for SignWriting.

Costa, A. C. R. & Dimuro, G. P. (2001) A SignWriting-Based Approach to Sign Language Processing. Escola de Informática, Universidade Católica de Pelotas

Sutton, V. (1999) Lessons in SignWriting – Textbook and Workbook. La Jolla: Deaf Action Committee for SignWriting (2nd ed.)

Sutton, V. & Gleaves, R. (1995) SignWriter – The world's first sign language processor. La Jolla: Deaf Action Committee for SignWriting.

Sun Microsystems, Inc. (2004) J2SE 1.5 „Tiger“ Feature List. Sun Microsystems, Inc. Santa Clara CA

Sun Microsystems, Inc. (2004) Java 2 SDK, Standard Edition, Version 1.5.0 - Summary of New Features and Enhancements. Sun Microsystems, Inc. Santa Clara CA

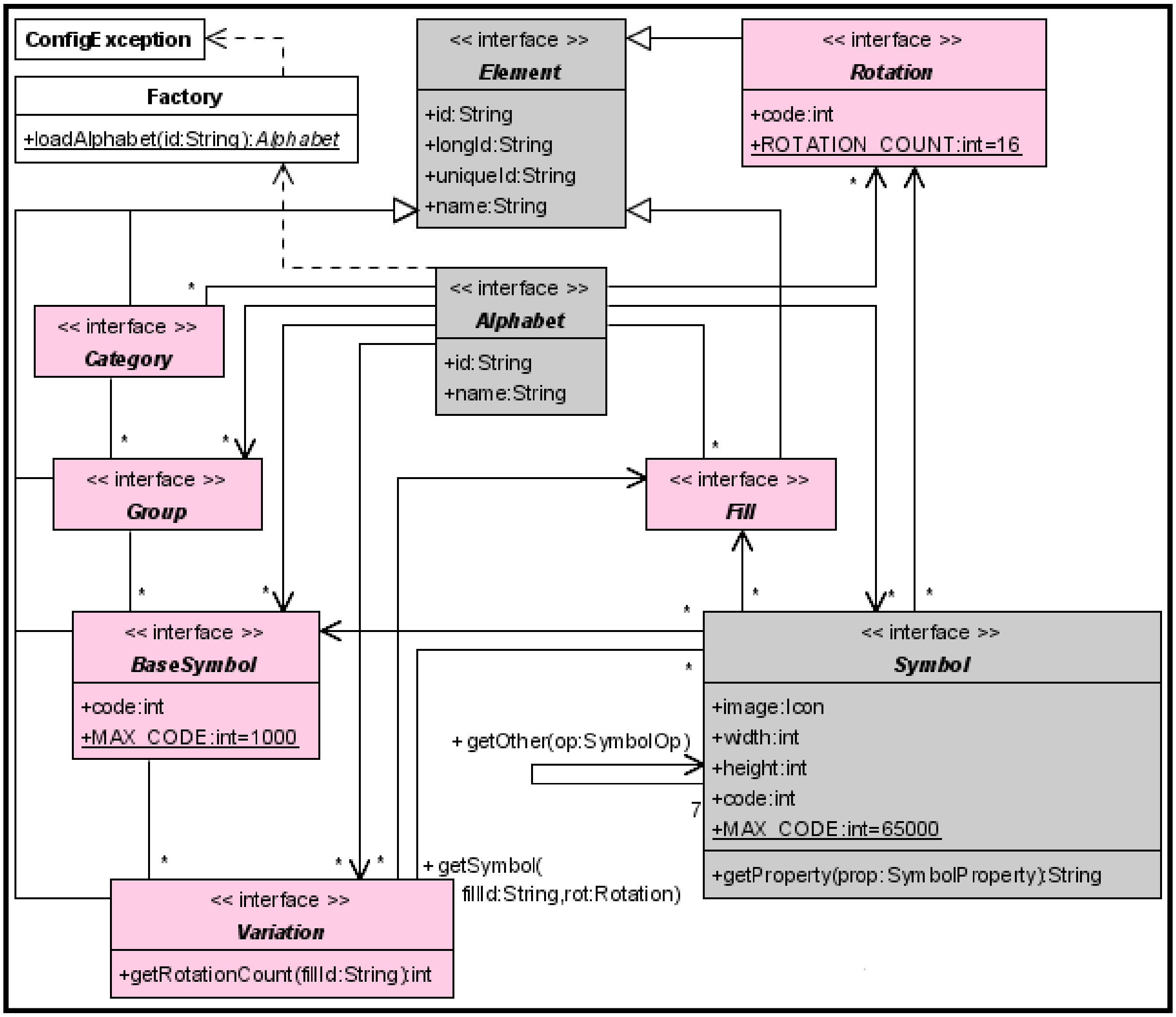


Fig. 1: UML class diagram for package signwriter.alphabet

An Overview of the *SiGML* Notation and *SiGMLSigning* Software System

Ralph Elliott, John Glauert, Vince Jennings, Richard Kennaway

School of Computing Sciences
University of East Anglia
Norwich NR4 7TJ, UK
{re,jrwg,vjj,jrk}@cmp.uea.ac.uk

Abstract

We present an overview of the *SiGML* notation, an XML application developed to support the definition of Sign Language sequences for performance by a computer-generated virtual human, or avatar. We also describe *SiGMLSigning*, a software framework which uses synthetic animation techniques to provide real-time animation of sign language sequences expressed in SiGML.

1. Introduction

We have developed the SiGML notation (Elliott et al., 2001) to support our work in the ViSiCAST and eSIGN projects (Glauert, 2002; Glauert et al., 2004). These projects have been concerned with the development of techniques for the generation of sign language performances by a computer-generated virtual human, or avatar.

The name SiGML is an abbreviation for "Signing Gesture Markup Language". SiGML is an XML application (Bray et al., 2004). Thus, SiGML data is represented as plain text in computer systems. SiGML encompasses several data formats used at different stages in the generation of virtual human animations, but its most prominent rôle is as the interface notation used in a prototype system supporting the generation of signed animation from natural language text. This system was a major part of the ViSiCAST project; as outlined in (Elliott et al., 2000), it contains two major subsystems:

- A "front-end" which uses natural language processing techniques to translate (English) text into an equivalent Sign Language form, for which a phonetic-level description is generated.
- A "back-end" which uses 3-D animation technology (together with artificial language processing) to generate a virtual human animation from the given phonetic-level description.

The natural language subsystem is designed to support output for several different national sign languages. Thus, it divides into a common initial stage, producing a language-neutral semantic representation (using DRT), followed by a stage specific to the target sign language. The most fully developed of the latter is that for British Sign Language (BSL) (BDA, 1992), which uses HPSG as the supporting grammatical formalism. More details on this work by our colleagues, Marshall and Safar, can be found in (Safar and Marshall, 2001; Safar and Marshall, 2002b; Safar and Marshall, 2002a; Safar and Marshall, 2002c).

The interface between the two subsystems is the SiGML notation, specifically the SiGML module we refer to as "gestural" SiGML. In the following section we describe gestural SiGML in more detail, concentrating on its relation to HamNoSys, the long established notation system for sign language transcription developed by our partners at the University of Hamburg. We then give a brief overview of SiGMLSigning, the back-end software subsystem identified above. We conclude with a simple example.

2. Gestural SiGML and HamNoSys

As we have indicated, gestural SiGML is based on HamNoSys (Prillwitz et al., 1989), that is, the Hamburg Notation System. This notation has been developed to support phonetic-level transcription of sign language performance by (real) human signers, and is intended to provide a model of sign language phonetics that is independent of any particular sign language. We have developed gestural SiGML with the explicit intention of formulating a model of signing gesture production which respects HamNoSys's model of sign language phonetics. At the start of the ViSiCAST project, HamNoSys stood at version 3. In preparation for the development of gestural SiGML, an initial phase of the ViSiCAST project saw the development of HamNoSys version 4 (Hanke et al., 2000; Hanke and Schmaling, 2002). As far as the manual aspects of signing are concerned, HamNoSys 4 does not radically alter the already well-established features of HamNoSys 3, but generalises and regularises several of those features. The more prominent changes in HamNoSys 4 occur in connection with the non-manual aspects of signing, for which a far more comprehensive framework is provided than was previously available. Following HamNoSys, gestural SiGML includes both a manual component, concerned with the configuration and actions of the hands, and a non-manual component, concerned with other linguistically significant features of signing such as head movement, eye movement, eye gaze, and mouthing. In the rest of this section we outline some general features of the SiGML notation before briefly describing the two components in turn.

2.1. General Features of Gestural SiGML

Considered as XML, a valid SiGML document is a pure element hierarchy: every element is constrained by the DTD (Kennaway et al., 2002) either to have element content or to be empty, that is, no SiGML element contains any embedded text, although of course it can, and in most cases does, contain attribute definitions. A SiGML document defines a sequence of "signing units". Typically, a signing unit is an explicit gestural definition for a single sign, but it may also be a direct definition of avatar animation parameters, or an indirect reference to another SiGML document. A gestural sign definition is represented by a <hamgestural_sign> element. Since it is intended that any HamNoSys sign definition can be represented in SiGML, we also allow a tokenised form of a HamNoSys sign, represented by a <hns_sign> element. For convenience of reference each of these sign elements has a `gloss` attribute, giving a (spoken language) gloss of the sign's meaning.

2.2. Manual SiGML

The manual component of a SiGML sign is represented by a <sign_manual> element. SiGML ascribes the same general structure to the manual component of a sign as does HamNoSys: an initial configuration followed by a sequence of actions or motions, which may well themselves be composite. Each of these components may involve both hands or just one hand, usually the signer's "dominant" hand (i.e. right hand for a right-handed signer). The initial configuration is a hand configuration, together with a location for that configuration. The configuration for each hand defines its hand shape, and its orientation in 3-D space. This orientation is specified as two components: extended finger direction (the direction of the metacarpal of the index finger) and palm orientation (the rotation of the palm about the axis defined by the other component). There is a basic set of a dozen standard handshapes, such as a fist, a flat hand, and a "cee" formed by the thumb and index finger. Many variations of these can be defined by specifying adjustments to the position of the thumb, various forms of bending of some or all fingers, and specific forms of contact or crossing between pairs of fingers. Hand shapes exemplify of HamNoSys's rather "operational" approach to the structure of feature definition: a simple instance of the given feature can be specified with no more than one or two symbols, while a more complex instance is obtained by appending additional modifier symbols defining how the required instance can be obtained from a simpler one.

In general terms, the location of a hand is defined with reference to a site on the signer's body, head, arm or (other) hand, and a rough measure of the proximity of the hand to that site. With some misgivings, we have retained in SiGML the HamNoSys concept of a "hand constellation", a special form of location which allows the definition of a potentially quite elaborate configuration of the hands as a pair, with (optionally) a location of this configuration relative to the body.

SiGML structures motions in a broadly similar fashion to HamNoSys, although SiGML tends to relegate to the level of informal semantics physical constraints to which HamNoSys gives direct syntactic embodiment. There is a repertoire of primitive motions, which may be combined in temporal sequence or in parallel, that is, concurrently, to any extent that makes physical sense. In SiGML, there are two other forms of structured motion (both inspired by comparable features in HamNoSys)

- Targeted motion: a motion for which an explicit target location (possibly a hand constellation) is specified.
- Repeated motion: various forms of single or multiple repetition of a given motion.

The simplest form of motion is a straight line motion in a given direction (any of the 26 directions defined by a non-zero position vector each of whose individual 3-D coordinates is either zero or one, or half-way between two adjacent directions of this kind). A straight line motion may be modified in a wide range of ways, including changing the distance moved, and tracing a curved, wavy or zig-zag path to the given end point. Other forms of simple motion include circular and elliptical motions (again with a wide range of variants), fluttering of the fingers, and several forms of wrist motion.

2.3. Non-Manual SiGML

The non-manual component of a SiGML sign is represented by a <sign_nonmanual> element. As described in (Elliott et al., 2004), the internal structure of this element closely follows non-manual feature definitions in HamNoSys 4. Thus, non-manual actions are partitioned into a hierarchy of tiers, corresponding to distinct articulators, as follows:

- Shoulder movements
- Body movements
- Head movements
- Eye gaze
- Facial expression: Eye-Brows, Eye-Lids, and Nose
- Mouthing: Mouth Pictures and Mouth Gestures.

Here, "facial expression" refers solely to those expressive uses of the face which are phonetically significant; by contrast those uses which express the signer's attitude or emotions about what is being articulated, important though they may be, cannot at present be expressed in SiGML (nor in HamNoSys). The two forms of mouthing reflect the distinction between motion of lips and tongue caused by spoken accompaniment to signing (mouth pictures), and other phonetically significant motions of lips, tongue, jaw and cheeks (mouth gestures). A mouth gesture often has a relatively elaborate internal structure which SiGML does not attempt to reflect, instead just identifying the unanalysed whole by a single label.

3. SiGMLSigning Animation Software System

SiGMLSigning is the software system we have developed, with support from partners in the ViSiCAST and eSIGN projects, to generate virtual-human signing animations on-screen from a sign sequence specified in SiGML. Architecturally, this system can be viewed as a pipeline of three processing stages, together with a control module which coordinates and schedules the transfer of data between these stages, stores the data they generate, and provides a programmable control interface. In its current form, the software is packaged as a set of Active X controls, which allow it to be deployed relatively easily in applications and

HTML pages on Microsoft Windows systems. The three processing stages are:

- SiGML Input and Pre-processing
- Animation Generation
- Virtual Human Animation

The interface between the first two stages is a sequence of gestural SiGML sign definitions; the interface between the second and third stages is a sequence of animation parameter sets, one set for each frame in the final animation. We outline each of these stages in turn, taking them in reverse order, in order to highlight the context each stage defines for its predecessor.

The final stage uses conventional 3-D animation technology. An avatar is represented by a virtual skeleton – a connected hierarchy of virtual bones – and a surface mesh – a connected tissue consisting of thousands of small, coloured, textured polygons. The configuration of these polygons determines the appearance of the avatar. The position and orientation of every polygon is determined (as part of the avatar's definition) by the position and orientation of one or more of the avatar's virtual bones. Hence a static posture of the avatar's surface appearance is completely determined by a static posture of its virtual skeleton: standard 3-D rendering techniques, using a combination of software and special-purpose graphics hardware, can be relied on to produce the one from the other. So, an animation of the avatar is defined simply by the appropriate sequence of static skeleton configurations, one for each animation frame (typically at the rate of 25 fps). A refinement of this system allows the avatar's appearance (in each frame) to be further modified by applying predefined distortions, known as morph targets or morphs, directly to the surface mesh. This technique is especially useful to us in defining facial non-manual gestures. The supplier of an avatar must therefore provide, as a minimum, a description of the physical structure of the avatar's skeleton and a list of its available morphs, together with a simple rendering interface which (i) allows a skeleton configuration to be specified (together with morph weights, if required), and (ii) accepts a request to render the corresponding posture.

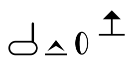
The preceding stage, at the heart of the SiGMLSigning system, is the animation generation stage, performed by a module called AnimGen. This maps a given sequence of gestural SiGML sign descriptions to the corresponding stream of avatar animation parameters. This stream is avatar-specific, since it depends crucially on the definition of the avatar's physical characteristics provided by the avatar supplier. Indeed, we have found that avatar-independent sign synthesis depends crucially on the specification by the avatar supplier of of the locations (relative to the skeleton) of quite a large number of sites on the avatar's surface mesh, in addition to the basic physical characteristics already mentioned. The task of this stage, therefore, is to derive precise numerical animation parameters from the physically relatively imprecise SiGML sign definitions. The manner in which this is done currently, and some of the issues that arise, have been described more fully elsewhere (Kennaway, 2001; Kennaway, 2003; Elliott et al., 2004).

The first processing stage performs relatively straightforward pre-processing of the SiGML input. Its most basic function is to decompose this input into individual sign definitions, so that each can be handled in the appropriate manner: the <hamgestural_sign>s can be fed directly to the AnimGen stage, the <hns_sign>s are first passed through a HamNoSys-to-(gestural-)SiGML translator, while those containing pre-generated animation data are just converted directly to the internal stored format output by the AnimGen stage, which is by-passed in this case. The HamNoSys-to-SiGML translation takes the form of an additional processing pipeline: conventional context-free parsing techniques (augmented with backtracking to account for HamNoSys's many syntactic ambiguities) are used to generate a syntax tree, which is then transcribed into an intermediate XML form, called HamNoSysML or HML; gestural SiGML is then generated from this using an XSLT transform (Clark, 1999; Kay, 2000).

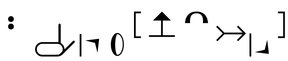
The SiGMLSigning software system is thus a "scriptable", virtual human signing animation system, accepting as input arbitrary signing sequences expressed in SiGML, and providing the corresponding animation on any avatar which supports the simple rendering interface described above. Finally, it is noteworthy that the core animation module, AnimGen, generates frames at a sufficiently high rate that the animation appears almost instantaneously in response to the SiGML input.

4. A Simple Example

The following is the HamNoSys sequence for a very simple gesture (which does not represent any actual sign):



Here, the first symbol specifies the hand shape, a fist with the index finger extended, the second and third symbols specify the orientation of the hand: the index finger points outwards from the signer's body, with the palm facing to the left; no initial location is explicitly specified for the hand, so a default, neutral, position in front of the signer's body is assumed; the final symbol specifies a straight movement from this initial position in an outwards direction, that is, away from the signer's body. The insertion of a few more symbols into this example results in a genuine sign, namely the DGS (German Sign Language) sign "going-to":



Here, the hand shape has a modifier specifying that the thumb is extended, the initial finger direction is now upwards-and-outwards, the outward motion has an upward arc modifier attached to it, and this motion is composed in parallel with a change of finger direction to downwards-and-outwards. The whole is prefixed with a symbol specifying motion of both hands in parallel, with the initial configuration of the non-dominant hand mirroring that of the explicitly specified dominant hand. The HNS-SiGML form of this is:

```
<?xml version="1.0" encoding="iso-8859-1"?>
<!DOCTYPE sigml SYSTEM .../sigml.dtd>
<sigml>
<hns_sign gloss="DGS_going-to">
  <hamnosys_manual>
    <hamsymmpar/>
    <hamfinger2/>
    <hamthumboutmod/>
    <hamextfingeruo/>
    <hampalml/>
    <hamparbegin/>
    <hammoveo/>
    <hamarcu/>
    <hamreplace/>
    <hamextfingerdo/>
    <hamparend/>
  </hamnosys_manual>
</hns_sign>
</sigml>
```

This is parsed during the input/pre-processing stage into the intermediate HML form shown (at the end of the paper) in Figure 2. In this easily generated but rather verbose format, an element typically corresponds to a HamNoSys syntactic category, while an attribute typically corresponds to an individual HamNosys symbol, although the HamNoSys parallel composition brackets and the HML `<paraction1>` elements provide a counter-example to this general rule of thumb.

The XSLT translation which is applied to the HML form shown in Figure 2 produces the much flatter Gestural SiGML form shown immediately below:

```
<sigml>
<hamgestural_sign gloss="DGS_going-to">
<sign_manual both_hands="true">
  <handconfig handshape="finger2"
    thumbpos="out"/>
  <handconfig extfidir="uo"/>
  <handconfig palmor="l"/>
  <par_motion>
    <directedmotion direction="o"
      curve="u"/>
    <tgt_motion>
      <changeposture/>
      <handconfig extfidir="do"/>
    </tgt_motion>
  </par_motion>
</sign_manual>
</hamgestural_sign>
</sigml>
```

The synthetic animation module, AnimGen, pre-processes this Gestural SiGML into a more explicit form of SiGML in which the hand-shape information is reduced to numerical measures of joint angles (on a scale of 1 to 4), and the rôle of both hands is made explicit. This explicit form is shown (at the end of the paper) in Figure 3.

The stream of animation data output by AnimGen is extremely voluminous, and is usually passed directly from the computer system's internal memory to the avatar rendering module. However, if desired, this data stream may be recorded for future reference in a file, in which case it is stored in SiGML's CAS (Character Animation Stream) format. A few lines of the output for our "going-to" example on the VGuido avatar, developed by our eSIGN project partner Televirtual, is shown in Figure 4 below.

The animation generated for this sign in isolation has a duration of about 320ms (preceded by another 320ms while the avatar's hands move from the rest position to the initial position of the sign itself. In Figure 4. below we show the animation frames for the start and finish of this sign.

Acknowledgements

We acknowledge with thanks financial support from the European Union, and assistance from our partners in the ViSiCAST and eSIGN projects.

5. References

BDA, 1992. *Dictionary of British Sign Language*. Faber and Faber.

Bray, T., J. Paoli, C.M. Sperberg, E. Mahler, and F. Yergeau (eds.), 2004. Extensible Markup Language (XML) 1.0. http://www.w3.org/TR/REC-xml/.

Clark, J., 1999. XSL Transformations (XSLT) Version 1.0. http://www.w3.org/TR/xslt.

Elliott, R., J.R.W. Glauert, and J.R. Kennaway, 2004. A framework for non-manual gestures in a synthetic signing system. In *Proc. Cambridge Workshop Series on Universal Access and Assistive Technology (CWUAAT)*.

Elliott, R., J.R.W. Glauert, J.R. Kennaway, and I. Marshall, 2000. Development of language processing support for the ViSiCAST project. In *ASSETS 2000 4th International ACM SIGCAPH Conference on Assistive Technologies*. Washington DC, USA.

Elliott, R., J.R.W. Glauert, J.R. Kennaway, and K.J. Parsons, 2001. D5-2: SiGML definition. Working document, ViSiCAST Project.

Glauert, J.R.W., 2002. ViSiCAST: Sign language using virtual humans. In *International Conference on Assistive Technology ICAT 2002*. Derby: BCS.

Glauert, J.R.W., J.R. Kennaway, R. Elliott, and B-J. Theobald, 2004. Virtual human signing as expressive animation. In *Proc. AISB-2004: Symposium on Language, Speech and Gesture for Expressive Characters*.

Hanke, T., G. Langer, C. Metzger, and C. Schmaling, 2000. D5-1: Interface definitions. working document, ViSiCAST Project.

Hanke, T. and C. Schmaling, 2002. HamNoSys 4 (Course Notes). http://www.sign-lang.uni-hamburg.de/Projekte/HamNoSys/HNS4.0/HNS4.0eng/Contents.html.

Kay, M., 2000. *XSLT – Programmer's Reference*. Wrox Press Ltd.

Kennaway, J.R., 2001. Synthetic animation of deaf signing gestures. In *4th International Workshop on Gesture and Sign Language Based Human-Computer Interaction*, LNAI. Springer-Verlag.

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*, LNAI, to appear. Springer-Verlag.

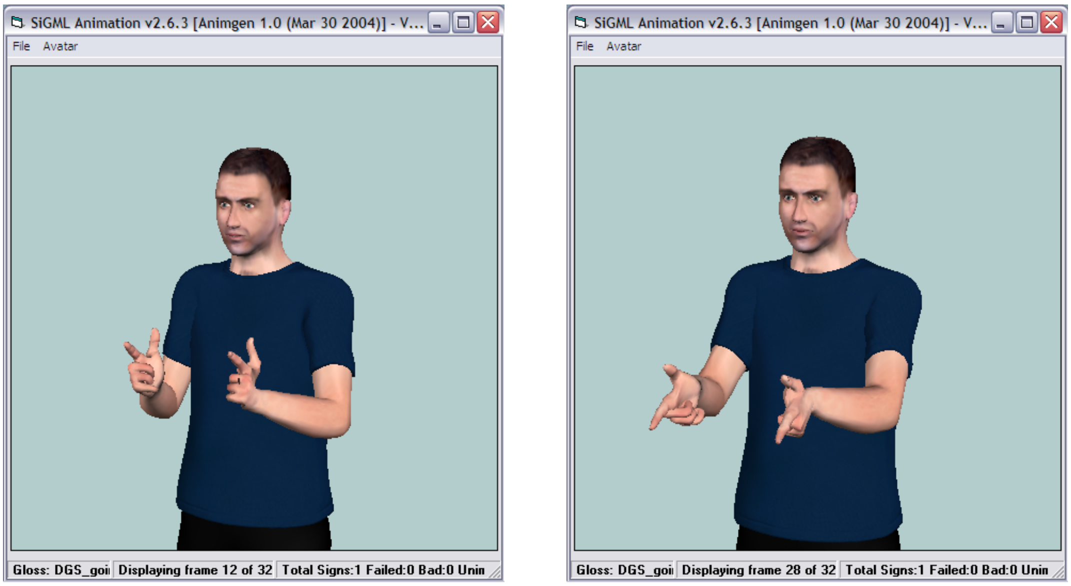


Figure 1: Animation frames for the “Going-To” Example.

Kennaway, J.R., R. Elliott, J.R.W. Glauert, and KJ Parsons, 2002. SiGML Document Type Definition (DTD). http://www.visicast.cmp.uea.ac.uk/sigml/sigml.dtd.

Prillwitz, S., R. Leven, H. Zienert, T. Hanke, J. Henning, et al., 1989. *HamNoSys Version 2: Hamburg Notation System for Sign Languages — An Introductory Guide*, volume 5 of *International Studies on Sign Language and the Communication of the Deaf*. SIGNUM Press, Hamburg.

Safar, E. and I. Marshall, 2001. Translation of English Text to a DRS-based, Sign Language Oriented Semantic Representation. In *Conference sur le Traitment Automatique des Langues Naturelles (TALN)*, volume 2.

Safar, E. and I. Marshall, 2002a. An Intermediate Semantic Representation Extracted from English Text For Sign Language Generation. In *7th Symposium on Logic and Language*.

Safar, E. and I. Marshall, 2002b. Sign Language Synthesis using HPSG. In *9th International Conference on Theoretical and Methodological Issues in Machine Translation (TMI)*.

Safar, E. and I. Marshall, 2002c. Sign language translation via DRT and HPSG. In *3rd International Conference on Intelligent Text Processing and Computational Linguistics (CICLing)*.

```
<hamnosysml>
<sign gloss="DGS_going-to">
<hamnosys_sign>
<sign2>
  <symmoperator att_par_or_lr="hamsymmpar"/>
  <minitialconfig2>
    <handconfig2>
      <handshape2>
        <handshape1 handshapeclass="ham_finger2" thumbpos="ham_thumb_out"/>
      </handshape2>
      <extfidir2>
        <extfidir1 extfidir="direction_uo"/>
      </extfidir2>
      <palmor2>
        <palmor1 palmor="ham_palm_l"/>
      </palmor2>
    </handconfig2>
  </minitialconfig2>
  <action2t>
    <action1t>
      <action1>
        <par_action1>
          <action1>
            <simplemovement>
              <straightmovement
                arc="ham_arc_u" movement="ham_move_o"/>
              </simplemovement>
          </action1>
          <action1>
            <simplemovement>
              <replacement>
                <extfidir1
                  extfidir="direction_do"/>
              </replacement>
            </simplemovement>
          </action1>
        </par_action1>
      </action1>
    </action1t>
  </action2t>
</sign2>
</hamnosys_sign>
</sign>
</hamnosysml>
```

Figure 2: Intermediate HML form for the “Going-To” Example.

```
<sigml/>
<hamgestural_sign gloss="dgs_going-to">
  <sign_manual both_hands="true">
    <handconfig handshape="finger2" thumbpos="out"
        bend2="0.00 0.00 0.00 0.00"
        bend3="4.00 4.00 4.00 0.00"
        bend4="4.00 4.00 4.00 0.00"
        bend5="4.00 4.00 4.00 0.00"
        bend1="-0.30 2.20 2.20 0.30 0.00"   />
    <split_handconfig>
      <handconfig extfidir="uo" palmor="l"/>
      <handconfig extfidir="uo" palmor="r"/>
    </split_handconfig>
    <handconstellation contact="medium">
      <location location="palm" bodyside="nondom" contact="touch"/>
      <location location="palm" bodyside="dom"    contact="touch"/>
      <location location="chest" contact="medium"/>
    </handconstellation>
    <par_motion manner="targetted">
      <directedmotion manner="targetted" direction="o" size="medium"
            curve="u" curve_size="medium" ellipse_direction="l"/>
      <tgt_motion manner="targetted">
        <split_handconfig>
          <handconfig extfidir="do"/>
          <handconfig extfidir="do"/>
        </split_handconfig>
        <handconstellation contact="medium">
          <location location="palm" bodyside="nondom" contact="touch"/>
          <location location="palm" bodyside="dom"    contact="touch"/>
        </handconstellation>
      </tgt_motion>
    </par_motion>
  </sign_manual>
</hamgestural_sign>
</sigml>
```

Figure 3: Explicit low-level SiGML for the “Going-To” Example.

```
<CAS Version="CAS2.0" Avatar="VGuido">
  <Frames Count="32">
    <Frame Duration="20.0000" BoneCount="67" MorphCount="42">
      <Morph Name="eee" Value="0.0000"/>
      ....
      <Bone Name="ROOT">
        <Position x="-0.0007" y="-0.0501" z="-0.0496"/>
        <QRotation x="-0.0286" y="-0.7137" z="0.0276" w="0.6993"/>
      </Bone>
      ....
    </Frame>
    ....
  </Frames>
</CAS>
```

Figure 4: Character Animation Stream (CAS) Data for the “Going-To” Example.

Statistical Sign Language Translation

Jan Bungeroth, Hermann Ney

Lehrstuhl für Informatik VI, Computer Science Department
RWTH Aachen University, D-52056 Aachen, Germany
{bungeroth, ney}@informatik.rwth-aachen.de

Abstract

In the field of machine translation, significant progress has been made by using statistical methods. In this paper we suggest a statistical machine translation system for Sign Language and written language, especially for the language pair German Sign Language (DGS) and German. After introducing the system's architecture, statistical machine translation in general and notation systems for Sign Language, the corpus processing is scetched. Finally, preliminary translation results are presented.

1. Introduction

The current progress in statistical machine translation suggests the usage of these methods on automatic Sign Language translation. This paper presents a first approach to such an application and discusses the advantages and disadvantages.

Deaf people, while fluent in their local Sign Language, often experience comprehension problems when they read written text or even lip-read spoken language. Thus for assisting the Deaf to communicate in a world of spoken languages, translation is needed. Currently human interpreters fill this gap, but their service is expensive and not always available. While a machine translation system can not fully replace an interpreter, it offers instant help in the everyday communication.

We therefore propose a system for translating a Sign Language into a spoken language and vice versa. Such a complete system translating from Sign Language to spoken language needs a gesture recognizer as input, the translation system and a speech synthesizer as output. The complete system translating from spoken language to Sign Language needs a speech recognizer as input, the translation system and a graphical avatar as output. In this paper the focus is held on the translation part. Figure 1 presents a schematic overview of such a system.

2. Related Work

In the recent years several groups showed interest in machine translation for Sign Languages.

- In our group, Bauer et al. (1999) proposed a framework for statistical-based Sign Language translation. The authors suggested to translate recognized video-based continuous Sign Language to spoken language.
- Other recent work was done by Sáfár and Marshall (2002) for translating English into British Sign Language using a rule-based approach. Here the grammar was modeled utilizing the HPSG formalism. The system is able to translate simple sentences.
- Huenerfauth (2004) introduces a rule-based concept for translating English text to American Sign Language (ASL).

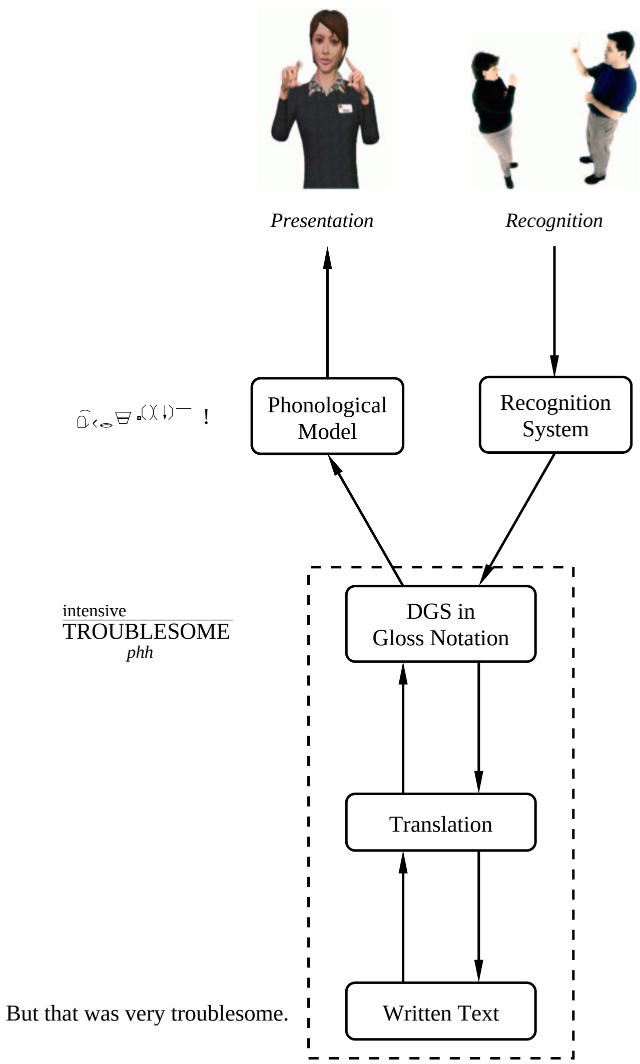


Figure 1: Automatic Sign Language translation system

- Also van Zijl and Barker (2003) propose another rule-based concept for translating English text to South African Sign Language (SASL).

Huenerfauth argues that a rule-based approach is better suited for Sign Language translation than statistical models because large corpora are difficult to obtain. He concludes that the use of a rule-based approach is more appropriate than the statistical. For our work, we do not think of this as an alternate option. Distinct corpora for Sign Languages are planned and already worked on. Additionally the optimization of the statistical translation process for scarce resources as suggested e.g. by Nießen and Ney (2000) allows for further improvement.

3. Statistical Machine Translation

Until recently, only rule-based systems were used for natural language translation. Such systems typically require hand written rules and dictionaries. However, over the last ten years a new approach has evolved, namely the statistical approach. This approach makes use of statistical decision theory and statistical learning. Such a system is trained using a set of sentence pairs. In recent evaluations like Chinese to English[1] and Arabian to English translations, it was found that these statistical approaches were comparable or superior to conventional systems.

In statistical machine translation a source sentence $f_1^J = f_1 \dots f_J$ is transformed into a target sentence $e_1^I = e_1 \dots e_I$ by choosing the sentence with the highest probability from all possible target sentences. This is given by Bayes' decision rule

$$\hat{e}_1^I = \underset{e_1^I}{\operatorname{argmax}} \{Pr(e_1^I) \cdot Pr(f_1^J | e_1^I)\}.$$

Several statistical models are used to estimate the free parameters with large training data (e.g. see Brown et al. (1993), Och and Ney (2000)). One target source word position is assigned to each source word position by alignments.

Figure 2 shows the general architecture of the statistical translation approach.

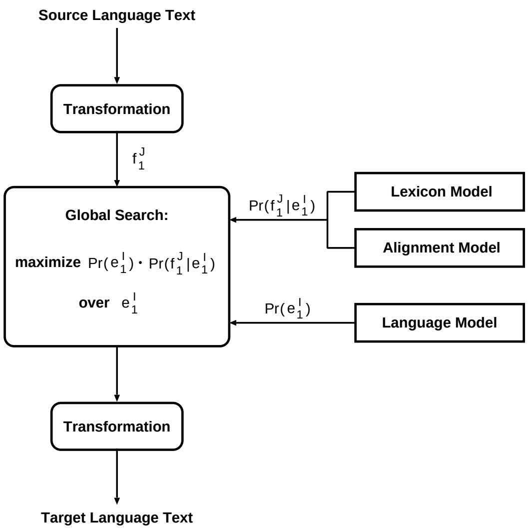


Figure 2: Architecture of the translation approach based on Bayes' decision rule

4. Notation Systems

Several different notations and phonological systems are common in Sign Language research. When dealing with Sign Language translation, an appropriate Sign Language representation is necessary to transfer data from and to the sign recognizer and the presentation avatar. Furthermore a word or phoneme based notation is needed for the internal alignment with the written words of the spoken language. A corpus based on such a notation system should qualify for learning and testing a statistical machine translation, but it might need pre- or postprocessing.

The following notation systems are introduced:

- Glosses are written words, where one gloss represents one sign. Additional markings provide further information, e.g. non-manual signs. Unfortunately no gloss standard exists, which results in inconsistent annotated corpora.
- The notation system introduced by Stokoe (1960) was the very first phonological symbol system of ASL. It divides signs into movement (sig), hand shape (dez) and location (tab) which occur simultaneously. As it focuses on ASL the application on other Sign Languages is not always possible. An ASCII encoding of the Stokoe system is available[2].
- The Hamburg Notation System HamNoSys (Prillwitz, 1989) is a more general form of the Stokoe system. Figure 3 shows an English sentence in gloss notation with markings and the corresponding HamNoSys glyphs.
- Liddell and Johnson (Liddell, 1984) suggest a sequential division of the sign stream into movement and hold segments. This avoids the simultaneous occurrence of phonemes.

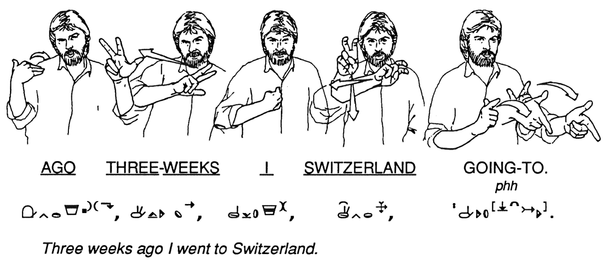


Figure 3: Example for HamNoSys and gloss notation taken from Prillwitz (1989)

5. Corpus Preparation

Statistical machine translation systems are trained using bilingual corpora containing full sentences. But two major problems arise when dealing with Sign Language. The first problem is the lack of large corpora. For example in written language, the Hansards corpus with French and English sentences from debates of the Canadian Parliament contains about 1,470,000 sentences. For Sign Language we have not found a corpus with more than 2000 sentences. The second problem is the lack of a notation standard. The

[1] http://nist.gov/speech/tests/mt/

[2] http://world.std.com/˜mam/ASCII-Stokoe.html

existent corpora use gloss notations which are too difficult to learn with limited corpora. Furthermore inconsistent use of the notation system complicates the problem.

For a starting basis the corpus collected by the DESIRE Team Aachen[3] consisting of 1399 sentences in DGS and German was investigated as it was one of the biggest available for us. Table 1 shows the details of this corpus, where singletons are words occurring only once. Note the very high number of singletons. This comes from the high diversity of the sentences. In addition, every word with a non-manual sign, e.g. (1), is counted as an extra word.

(1) $\overline{\text{HABEN}}^{\text{neg}}$
"not have"

	DGS	German
no. of sentence pairs	1399	
no. of running words	5480	8888
no. of distinct words	2531	2081
no. of singleton words	1887	1379

Table 1: DESIRE corpus statistics

This is not usable for statistical machine translation. Thus for first experiments a small corpus was built from the DESIRE corpus. Several considerations were made:

Brackets indicating a non-manual sign on a whole phrase or sentence are expanded. Consider the sentence (2).

(2) WAHL+ERGEBNIS $\overline{\text{WISSEN DU}}^{\text{qu}}$
"Do you know the election results?"

Table 2 shows the ASCII representation of this sentence before and after expanding the brackets.

```
WAHL+ERGEBNIS qu-{WISSEN DU}
WAHL+ERGEBNIS qu-WISSEN qu-DU
```

Table 2: Expanding brackets in the corpus file

Additional information to locus agreement was deleted as it can not be learned. E.g. in the phrase (3) the 'arbeit' refers to a place in signing space. This information is deleted. After the translation to DGS it can be partially reconstructed by rules.

(3) ARBEITEN X'arbeit'
"at work"

When suitable, the non-manual signs were treated as single words. As an example (4) is processed as seen in table 3, so it can be mapped to the German translation "nicht mögen". But (5) is kept so it can be mapped to the German "unmöglich".

(4) $\overline{\text{MÖGEN}}^{\text{neg}}$
"to like not"

(5) $\overline{\text{MÖGLICH}}^{\text{neg}}$
"impossible"

[3]http://www.germanistik.rwth-aachen.de/desire

```
neg-MÖGEN
neg MÖGEN
```

Table 3: Separating non-manual signs in the corpus file

These methods were used to form the new corpus of 200 sentences. In this corpus the number of singletons is kept low for better training. In addition most words or word forms have an entry in a bilingual manual lexicon. Table 4 gives an overview of the corpus. While this is not enough training data for a fully-fledged translation system, it allows the first experiments, we will discuss in section 6.

		DGS	German
Training:	no. of sentence pairs	167	
	no. of running words	845	828
	no. of distinct words	73	142
	no. of singleton words	15	48
Testing:	no. of sentence pairs	33	
	no. of running words	157	161
	no. of distinct words	43	74
	no. of singleton words	18	40

Table 4: The small DGS/German corpus statistics

6. Results

For translation experiments, training and testing data is needed, as-well as an objective error measurement. The corpus shown in table 4 is divided into training samples (83% of the sentences) and testing samples (17% of the sentences). The training is performed by using various statistical models like IBM Model 1-4 (Brown et al., 1993) and others like Hidden Markov Models HMM (Och and Ney, 2000). Figure 4 shows the alignment of a sentence pair which is obtained in training. For testing, the test sentences

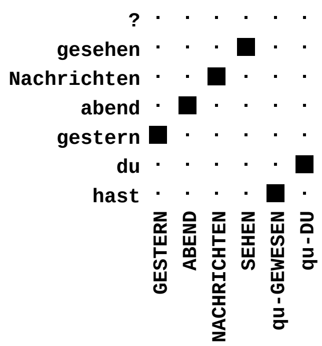


Figure 4: Trained alignment of a sentence pair

in the source language are translated and compared with the the known target sentences. These translation results are evaluated.

We use the following objective evaluation criteria for error measurement:

German	automatic DGS translation	manual DGS translation
du wartest darauf daß der Tee kommt	DU WARTEN BIS TEE KOMMEN	DU WARTEN BIS TEE KOMMEN
frische Bananen und Äpfel schmecken gut	FRISCH ÄPFEL UND BANANEN SCHMECKEN GUT	BANANEN FRISCH UND ÄPFEL SCHMECKEN GUT
ich mag nicht fliegen	ICH NICHT UNKNOWN_fliegen	FLIEGEN ICH neg MÖGEN

Table 5: Translated sentence pairs for German and DGS

- mWER:
  The word error rate (WER) is computed as the minimum number of substitution, insertion and deletion operations that have to be performed to convert the generated sentence into the target sentence. This performance criterion is widely used in speech recognition. This minimum is computed using a dynamic programming algorithm and is typically referred to as edit or Levenshtein distance. In addition for the multi-reference WER (mWER) not only one but a set of reference translation sentences is used. (Nießen et al., 2000)

- mPER:
  The position-independent word error rate (PER) compares the words of the two sentences without considering the word order. The PER is less than or equal to the WER. The multi-reference PER (mPER) again considers a set of reference translation sentences.

We performed the translation from German to DGS on the small corpus. Table 6 shows the mWER and mPER error rates for our experiments. As a reference the baseline is a single word-to-word translation. We then applied our models for the training of alignment models to improve the results.

	mWER [%]	mPER [%]
single word	85.4	43.9
alignment templates	59.9	23.6

Table 6: Testing results for German to DGS

The examples in table 5 show translations from our test corpus. The first sentence is a correct translation, while the second sentence is in partial disorder. The last sentence shows a wrong word order and missing words.

7. Summary

For the translation of spoken language into Sign Language, we propose statistical machine translation. Such a system is trained with bilingual corpora. While Sign Language corpora are still rare, we demonstrated how such a corpus can be prepared for the translation system. Furthermore we performed first experiments on a small German-DGS corpus and presented results. While this is meant only as a small-scale example and a proof-of-concept, we are confident of applying our methods to real-world conditions and corpora.

Future work includes the construction of a more suitable corpus and further improvement of the translation performance. Especially we expect performance gain from the use of better dictionaries and linguistic knowledge like morpho-syntactic information.

8. References

B. Bauer, S. Nießen, and H. Hienz. 1999. Towards an automatic Sign Language translation system. In *Proc. of the Int. Workshop on Physicality and Tangibility in Interaction*, Siena, Italy.

P. F. Brown, S. A. Della Pietra, M. J. Della Pietra, and R. L. Mercer. 1993. Mathematics of statistical machine translation: Parameter estimation. *Computational Linguistics*, 19(2):263–311.

M. Huenerfauth. 2004. A multi-path architecture for machine translation of English text into American Sign language animation. In *Proc. Student Workshop at Human Language Technolgies conference HLT-NAACL*, Boston, MA, USA.

S. Liddell. 1984. Think and believe: Sequentiality in American Sign Language. *Language*, 60(2):372–399.

S. Nießen and H. Ney. 2000. Improving SMT quality with morpho-syntactic analysis. In *Proc. on the 18th Int. Conf. on Computational Linguistics*, Saarbrücken, Germany.

S. Nießen, F.J. Och, G. Leusch, and H. Ney. 2000. An evaluation tool for machine translation: Fast evaluation for machine translation research. In *Proc. of the Second Int. Conf. on Language Resources and Evaluation (LREC)*, pages 39–45, Athens, Greece.

F. J. Och and H. Ney. 2000. A comparison of alignment models for statistical machine translation. In *Proc. on the 18th Int. Conf. on Computational Linguistics*, pages 1086–1090, Saarbrücken, Germany.

S. Prillwitz. 1989. *HamNoSys. Version 2.0; Hamburg Notation System for Sign Language. An Introductionary Guide*. Signum Verlag.

E. Sáfár and I. Marshall. 2002. Sign Language generation using HPSG. In *Proc. Int. Conf. on Theoretical and Methodological Issues in Machine Translation*, pages 105–114, TMI Japan.

W. Stokoe. 1960. Sign Language structure: An outline of the visual communication systems of the American deaf. *Studies in Linguistics, Occasional Papers*, 8.

L. van Zijl and D. Barker. 2003. South African Sign Language machine translation system. In *Proc. 2nd Int. Conf. on Computer graphics, virtual Reality, visualisation and interaction in Africa*, pages 49–52, Cape Town, South Africa.

Computer support for SignWriting written form of sign language

Guylhem Aznar, Patrice Dalle

TCI team, IRIT lab www.irit.fr
<lrec@externe.net>, <dalle@irit.fr>

Abstract

Signwriting's thesaurus is very large. It consists of 425 basic symbols, split in 60 groups from 10 categories. Each basic symbol can have 4 different representations, 6 different fillings and 16 different spatial rotations.

While signwriting is more and more used by the deaf community, it currently lacks a complete and platform neutral computer support to let signwriters share documents regardless the applications and the underlying operating system they may be using.
Based on previous research, various propositions have been made, resulting in multiple incompatible systems.
The main problem currently is the lack of a consistent basis upon which compatibility could be built : the most advanced and used system, SWML [1], is multiplatform thanks to Java but requires dedicated applications like the previous attempts.
Moreover, the use of XML based representation requires dozens of lines of code for each symbol, resulting in oversized files which can not be parsed, used or read with standard tools. XML linking to bitmap pictures for on-screen representation prevents the integration of a real font system, needed for a true portability, and cause scalability problems.
Moreover, like previous systems, SWML still comes with a complex user interface, a little easier to learn but slower, symbols being entered via the mouse.
Even if this advanced approach helped the signwriter community, replacing the manual insertion of GIF graphic files for each symbol, at the moment, the signwriting community must revert to screenshots and pictures to ensure documents can be shared and read, resulting in little reusability for both users and researchers, and low computational possibilities worsened by the absence of signwriting optical recognition software. Guylhem Aznar, a first year medical resident and a PhD student in Computer Science from Pr. Patrice Dalle TCI team in IRIT (Toulouse, France), is proposing a unicode based representation for Signwriting with a suite of free software tools running on GNU/Linux but also supporting non-free operating systems.
This approach based on unicode is putting a strong emphasis on facilitating communication and compatibility through a unicode reconstruction engine.
Usage and computer entry are also made simpler thanks to different possibilities of human interaction : keyboard, mouse and sensitive area (handwriting) support, which all result in the same unicode-text output. This output can then be shared, reused or studied easily.
The choice of unicode over XML facilitates integration in existing software. The system works in layers : the entry layer, the keycode layer, the unicode layer, the rendering layer and the font layer. These layers are independent and therefore easy to adapt and improve. In the keycode layer, each signwriting “basic symbol” is coded by a different number called “internal name”.
This basic symbol is first positionned geometrically by “positionning elements” defining concentric circles and the respective angular position of the basic symbol on these circles. The basic symbols can be completed by additional information regarding the possible variations, such as spatial rotations, required in order to form the complete symbol. These “additional information elements”, like the basic symbols and the positionning elements, are also coded by one or more numbers also called internal names.
All these internal names are linked to their respective meanings in a mapping table. Additional internal names can be defined following the evolution of signwriting's standard. Finally, “delimitors” are used to group basic symbols into complete signwriting units. In the unicode layer, another mapping table is used : these internal names are mapped to unique unicode characters. One or more internal name can be mapped to a unicode character, but each unicode character can only have one mapping. This non-bijective approach is required to follow the unicode standard.
In the entry layer, signwriting symbols can be entered by different peripherals like a keyboard or a mouse. The mouse driven graphical input system will be completed by other entry modes in the future. Following the traditional key mapping entry mode, a table maps internal names to the physical keys on the keyboard. Multiple keyboard mapping tables allow different physical dispositions for different countries or following user preferences.
The entry layer is separated from the rest of the system. It is only relevant to the system by its dependancy on the unicode layer, required in order to output Unicode characters following the keycode layer specifications.
In the rendering layer, a unicode reconstruction engine like Gnome's Pango, transform the flow of unicode characters into a graphical representation, i.e. a complete signwriting symbol. It is not yet suitable to the display: elements are still numbers (then called *“external names”*), and must be replaced by graphics.
The transformation is coded by a set of rules [3] describing the possible combination and the outputs, like for unicode arabic and indian languages support. In the font layer, a font subsystem like Gnome's Freetype/xft2,

which support both traditional bitmap fonts and vectorial fonts, takes care of the graphical representation, replacing external names by their corresponding graphical symbols. Different fonts can of course be used.

Considering a symbol has been entered though the entry layer, it must then be transcribed into a serie of unicode characters following these steps:

- first, a positionning element is used to define a circle. If this circle preceded by another circle before the initial delimitor, it is embedded in that circle. A special type of circle is used to define the contour of the face
- then, basic symbols are positionned on that circle, with positionning elements to define their angular position followed by additional information elements if these basic symbols need rotations, special fillings, etc.
- finally, a delimitor is used to mark the end of the signwriting unit.

The internal names of these entities are never used – instead, unicode characters are used, which allows existing software to process signwriting. These Unicode caracters are then mapped to the internal names, and the rendering layer geometrically and spatially reconstruct a complete signwriting unit in the form of external names. The font layer then replaces this information by the graphical drawing of the complete unit.

Currently, the different layers are under work. They do not require the same amout of work: the most complicated part is the definition of rules for the rendering layer [4], the hardest task is drawing fonts, the most important is the keycode layer to provide a quick replacement to SWML and the longest part is reserving enough space in unicode for a correct signwriting implementation. The latter may eventually be impossible, in which case "private" unicodes areas will have to be used. This should only cause some minor changes in the unicode layer, but will damage the portability benefits of using unicode.

This entire "text-like" layered approach makes a clear separation between the various sub-systems used, providing a solid base upon which new sub-systems can be built (for ex. in the entry layer, handwriting recognition) and any layer can be upgraded (ex: adding additional vectorial fonts, supporting a new signwriting standard) without requiring a full system redesign. Applications following Gnome's API can immediately take advantage of signwriting support, which means a whole desktop suite of software is made available for free to deaf-users. Moreover, signwriting features (ex: writing from top to bottom) no longer need special handling through specific applications, thanks to Gnome localisation support.

An additional advantage is the portability of the model. Support on the GNU/Linux based PDAs requires no further work. Windows or MacOS support would require minimal support in the entry layer and at some specific points in the font layer.

The upcoming support of Windows and MacOS by Gnome applications means these steps could also simply be removed in the short term. Moreover, Signwriting transcription in standardized unicode text means the text can be subject to automated computer analysis, exchanged by researchers, etc. Possible evolutions of the system include a statistical approach for auto completion and handwriting recognition, and will certainly focus on the user interface with the design of specific Gnome Accessibility features.

1. References:

[1] Rosenberg, A. Writing Signed Language, "In Support of Adopting an ASL Writing System", Dept. of Linguistics, Univ. of Kansas, USA, 1999 http://www.signwriting.org/forums/research/rese010.html

[2] Antonio Carlos da Rocha Costa and Gracaliz Pereira Dimuro, "A SignWriting- Based Approach to Sign Language Processing", Universidade Catholica de Pelotas, Brasil, 2001 http://www.techfak.unibielefeld.de/ags/wbski/gw2001book/draftpapers/gw38.pdf

[3] Klaus Lagally, "ArabTe{X} : typesetting arabic with vowels and ligatures", 1992 http://citeseer.nj.nec.com/rd/64325440%2C95192%2C1%2C0.25% 2CDownload/http://citeseer.nj.nec.com/cache/papers/cs/350/ftp:zSzzSzftp.informatik.unistuttgart.dezSzpubzSzlibraryzSzncstrl.ustuttgart_fizSzTR-1992-07zSzTR-1992-07.pdf/lagally92arabtex.pdf

[4] Finite State Automata and Arabic Writing - Michel Fanton Certal-Inalco http://acl.ldc.upenn.edu/W/W98/W98-1004.pdf

Chinese Sign Language Synthesis and Its Applications

Yiqiang Chen, Wen Gao, Changshui Yang, Dalong Jiang, Cunbao Ge

Institute of computing technology, Chinese Academy of Sciences, Beijing, China, 100080
{yqchen, wgao, csyang, dljiang, cbge}@jdl.ac.cn

Abstract

The Sign Language is the communication language for deaf-mute community. Everywhere in the world may have their own sign language. There are over 20.57 million deaf people with thousand kinds of language in China. Hence a set of standard Chinese Sign language for the Deaf-Mute has been revised several times by Chinese Deaf-mute Association supported by the Chinese Government. The updated standard Chinese sign language will help you easily communicate with any deaf people in China.

"How can we learn it in a short time and convenience way"? The traditional face-to-face and tape recorder teaching methods can't express meaning well due to time and space limitation. Then a Chinese sign language synthesis system has been developed. The system uses artificial intelligence and computer graphic technology to demonstrate the Chinese sign language from optional customers by a 3-dimensions virtual human. The software will help to demonstrate a standard Chinese sign language by a 3-dimensions virtual human if you print Chinese language (Fig.1).

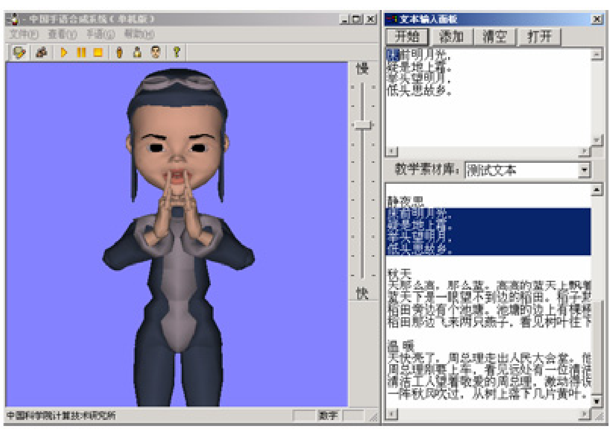


Fig.1: Chinese Sign Language Synthesis System

This system has integrated advanced research results and key technologies both domestic and abroad. There are three new technologies being developed in our system. First is realistic face animation, in sign language, there are about 1/3 words of the whole sign words set must have expressions to make the gesture understood. Therefore, face animation and expression synthesis is more important for the whole system. Second is motion retargeting technology, which can retarget the standard data to any given character modal to make the animation data singer-independent. And the third is synchronization modal between gesture and lip motion. There are also several characters of the system: first is the system covered large vocabularies of Chinese Sign Language, totaling up to 5596 sign language terms, 30 Chinese finger languages and 24817 synonym, almost contained all of contents of middle schools and elementary schools text book in China. Second is it realized study interaction and not limited by time and space. Third is customers could choose from several human images or models. And last are clear interface, easy operating and free adding new sign language. The experiment shows that the software was given a score of 92.98 for visual and understanding finger spelling, 88.98 for words, 87.58 for sentences by the students from these Deaf-mute Schools. The system has a great significant for standardizing, studying and popularizing Chinese Sign Language and plays a very important role for the hearing impaired society. This system could be used in all kinds of service business and public places, and could bring a great convenience for the deaf-mute's life and study, etc WebSigner (Fig.2(a)), TVSigner (Fig.2(b)), OnlineSigner (Fig.4(b)).

(a) (b) (c)

Fig.2: The system applications

The WebSigner (a) can be used for aiding deaf person to obtain information from Internet using his convention way. The TVSigner (b) can generate virtual signer for TV that aids deaf person to watch TV. The OnlineSigner (c) can be used for Chinese sign language learning, which allows you learn standard Chinese sign language in a short time.

“Progetto e-LIS@”

Paola Laterza and Claudio Baj

ALBA Piccola Soc. Coop. Sociale Via San Donato, 81 – 10144 TORINO (Italy) E-mail: coop.alba@mclink.it

"Progetto e-LIS@" is the presentation of a work-in-progress, which was started in November 2000 by two Italian scholars, Paola Laterza (who is a hearing psychologist) and Claudio Baj, a Deaf LIS teacher. Their aim is to find a system of cataloguing signs in order to create a complete but flexible multimedial dictionary, to be used both by competent Italian Sign Language users and by competent users of Italian. This research presents a new way of ordering signs, different from the usual alphabetical one, and is more congenial to the signing community's linguistic needs, which are clearly oriented to the visual-corporeal channel rather than to the written-oral one. In fact, there are Italian/Sign Language dictionaries based on the alphabetical order, but there is none that goes from Sign Language to the written-oral language (Italian). Special attention has been paid to how signs are systematised: so far the handshape parameter has been explored in detail, but in the near future we plan to associate it with two more parameters, viz. location and orientation. At a later date movement and non-manual signals will also be included among the cataloguing criteria. The objective is not only to put signs in order according to a more flexible and therefore acceptable system for signers (like the alphabetical order satisfies hearing people's phonological needs), but also to allow for the quick search of signs in the multimedial dictionary. The paper describes how, after elaborating different versions in their step-by-step research, the two researchers decided that the present format was more functional, practical and economical from the point of view of the dictionary as an instrument. They will present the results already obtained in their research as well as their intermediate findings to demonstrate their chosen work method but also to receive feedback from other Italian and European realities.

1. HANDSHAPES

1 st version (27th November, 2000)
Figure 1: 1st version
Our first step was to single out a number of so-called “principal” handshapes, chosen from the ones that appeared clearest, best-defined, with extended fingers and in alignment with the hand, easy to remember for either experienced or inexperienced signers. 14 handshapes were chosen: As – A – S – G – I – L – Y – H – V – Ycor – 3 – 4 – B – 5. These were ordered by starting from the closed fist and progressing to the open hand, since we recognized the fist as the origin of all the other handshapes (cf. Volterra 1987). Subsequently one finger at a time appears from the fist-shape: first the thumb, then the index finger and the numbers from zero to five. The thumb represents the number 1, the index 2, and so on, up to 5 with the open hand. Then there are also two fingers that appear simultaneously, then three, and so on, up to the point of having five extended fingers and an open hand. When there are two handshapes that have the same two extended fingers, preference is given to the one with two joined fingers rather than to the one with open fingers, because the latter looks more open from the visual level (e.g. “H” vs. “V”). In cataloguing the handshapes, reference is made to the dominant hand, even if a sign requires both hands with different handshapes. The handshape symbols have been taken from the dictionary by Radutzky (1992).

2 nd version (15th January, 2001)
Figure 2: 2nd version
In the second version we maintained the same criteria as in the first, but a few slight changes were made in the choice of the principal handshapes. We felt the need for a further criterion which would allow us to flexibly insert as many handshapes as possible by following an order that will not create confusion. Therefore we saw the addition of the subgroup criterion as a useful innovation. The principal handshapes are still 14, but with some variations. The sequence of the hanshapes were changed vis à vis the previous version. However, the number of principal handshapes remained unchanged. The new order of principal handshapes was as follows: A – S – G – I – L – Y – H – V – Ycor – 3 – 4 – 3/5 – B – 5. Over and above the 14 principal handshapes, we started to include other “subordinate” handshapes, putting them in subgroups dependent on the principal ones. The subgroups were catalogued according to the position of the fingers in the principal handshapes, from which, with progressive curving, bending or closing movements, one finally reached the subordinate subgroups. After singling out the subgroup criterion we chose to add a further criterion to order the handshapes within the subgroups themselves. According to this criterion, the subordinate handshapes follow a contrasting closing-opening movement, followed by the principal handshapes: starting from the maximum opening of the principal handshape, the subgroup is shaped by the progressive closing of the fingers (e.g. L, cc, Lp, Lq, Lch, T). In this version 37 handshapes were catalogued.

3 rd version A (12th March, 2001)

Figure 3: 3rd version A

Here we followed up our previous findings and tried to add more and more handshapes, but at the same time maintaining clarity and linearity. To facilitate our research for the multimedial dictionary, we decided to subdivide the subgroups further, creating branches of the principal handshapes. In the previous version each subgroup was linear and the handshapes (both curved and flat) were collocated within it and ordered according to a very arbitrary criterion of closure based on the impression of more or less filling of the visual space. Here, on the other hand, some branches were drawn up from those handshapes which, starting from the principal one, follow a movement of flat closure while other branches follow a movement of curved closure. This version includes 53 handshapes, of which 20 are principal, and represents an attempt to list and order all the handshapes existing, in our opinion, in Italian Sign Language. The principal handshapes are: As – A – S – G – I – L – Y – H – V – Ycor – Hs – 3 – Ys – W – 4str – 4 – 3/5 – B – Bs – 5. In this version we started to systematize the criteria; some remained unchanged while other new ones were created from the previous versions. First criterion: the order of the principal handshapes proceeds from the closed fist to the progressive extension of one finger at a time, from the thumb to the little finger, and subsequently of two, three, four and five fingers extended at the same time. In the first five handshapes, each finger is withdrawn to leave space to the following one, following the numerical order from the thumb to the little finger. The same principle guides the order of the handshapes formed either by pairs of fingers, by threes, fours or fives. Second criterion: among the principal handshapes, according to the principle of progressive opening of the hand, those with joined fingers precede those with the same but separate fingers. Third criterion: having chosen to consider all handshapes as independent/separate from each other, we decided that a linear, sequential list of 53 handshapes would be difficult to implement. To overcome the difficulties that a very long list would cause in cataloguing, in learning, memorizing and use, already during the second version we opted for the creation of subgroups. As "principal" handshapes we chose handshapes which were clearly contrasting with each other and easy to perform from the motorial point of view. The subgroups consisted of those "subordinate" handshapes that present limited distinctive features and are more difficult to perform. Fourth criterion: since the principal handshapes, chosen from the clearest and most distinct, are performed with the fingers in an extended position and in alignment with the hand, it is obvious that the movement necessary to order the subgroups follows the progressive closure of the fingers, contrary to the movement of progressive opening of the principal handshapes. Fifth criterion: since an enormous variety of subordinate handshapes exist within the subgroups, we have tried an ulterior subdivision to create more order. Different branches originate from a principal handshape, depending on the typology of the closure movement (i.e. flat or circular). The flat handshapes moving towards progressive closure with extended fingers precede the handshapes with curved fingers, since the latter enclose a more limited area of the palm, while the former leave a wider opening. Sixth criterion: in signs where both hands are used, handshapes are sometimes different. In cataloguing these cases, reference is made to the dominant hand (i.e. for right-handed people, the right hand, and for the left-handed, the left hand).

3 rd version B (6th February, 2002)

Figure 4: 3rd version B

In the following version the previously elaborated criteria have undergone some more changes; moreover, five new handshapes have been added to reach a total of 58. Some movements have also been carried out, to better satisfy the recognized criteria. We thus have 20 principal handshapes: S – G – Yi – I – L – Y – H – V – Ycor – Hs – 3 – Ys – Wstr – W – 4str – 4 – 3/5 - B – Bs – 5. The new handshapes are those that are used very little but are present in LIS and have never been catalogued officially. First criterion: The principal handshapes have fingers extending from the fist at a right angle and are not bent, while the other fingers are closed, i.e. they have contact with the palm of the hand. They have been singled out among those handshapes which correspond to the numbers "1" to "5" in one hand, starting from the thumb and ending at the little finger. Second criterion: The principal handshapes follow the movement of progressive extension of the hand from a closed to an open position, from "1" to "5", from the thumb to the little finger, following the intermediate passages. Third criterion: In the subordinate handshapes the fingers are in a bent position, and, if they have contact with parts of the hand, it is not with the palm (like in the closed handshapes), but almost exclusively with the fleshy tip of the thumb. Other contacts between fingers are only considered as part of movement of large closure. Subordinate handshapes are grouped together in subgroups. Fourth criterion: Since the fingers are straight and not bent in the principal handshapes, it is selfevident that the subordinate handshapes follow a movement of progressive closure within the subgroups, in contrast with the movement of progressive opening in the principal handshapes. Fifth criterion: Different branches originate within a subgroup from a principal handshape, depending on the typology of closure movement applied, i.e. flat or circular. Flat handshapes closing progressively with extended fingers precede the ones where the fingers are curved, since the latter occupy a more limited area of the palm, while the former allow for a larger opening. Sixth criterion: In signs where both hands are used (Volterra, 52), the handshapes are sometimes different. For purposes of cataloguing, in these cases reference is made to the dominant hand (i.e. the right hand for right-handed people and vice versa for the left-handed). For the purposes of this research, the latter version is presently considered the most functional, the clearest and simplest for ordering signs. The criteria that have been emphasized are definitive, in the present state-of-the-art.

2. COUNTERCHECKS

During the work-in-progress, when the criteria for cataloguing the signs had been established, we looked for counterarguments and confutations which could show which of these were fundamental, superfluous or arbitrary, but keeping version 3B as the reference point. In this way we started to build up new versions.

1 st countercheck (3rd July, 2002)
Inversion of criterion 2 Main handshapes Figure 5: 1st countercheck
The 21 principal handshapes have been put into a particular order by inverting criterion 2, i.e. from the maximum to the minimum opening of the fist closure, to see if this criterion is fundamental or arbitrary. Building up the scheme, this criterion proved to be arbitrary, since exclusion or confusion of handshapes does not result from the inversion of the order. This countercheck did not include subgroups. In the next counterchecks we shall see if the order remains functional when subgroups and other criteria are added.

2 nd countercheck (2nd October, 2002)
Confutation of criteria 1 and 3 Linear sequence
Figure 6: 2nd countercheck
The 58 handshapes have been ordered according to the progressive opening of the hand without creating subgroups (criterion 3), and therefore in a linear sequence. We saw that, in this way, groupings of handshapes according to finger positions did not take place if no distinction between principal and subordinate handshapes (criterion 1) was effected. The sequence of resulting handshapes was therefore determined randomly and exclusively through the perception of the hand more or less filling the visual space. Moreover, in a similar sequence, it was impossible to single out a simple logic to understand and memorize: remembering 58 elements without any clear, precise reference points proved to be difficult. Thus we concluded that is was necessary to single out principal handshapes and subgroups in order to produce an applicable order. Therefore criteria 1 and 3 proved to be fundamental. The order in which fingers open up could be inverted, from the open hand to a fist, but there were no structural changes and no handshapes were excluded. In this way the arbitrariness of criterion 2 was confirmed.

3 rd countercheck (9th April, 2004)
Inversion of criterion 2 With subgroups Figure 7: 3rd countercheck
Inverting the order of the principal handshapes and going against criterion 2, i.e. from the open hand to the closed fist, and following the creation of subgroups according to criteria 3, 4, 5 and 6, leads to the reproduction of a "version 3b" in reverse, but without making it less clear or organized. In this way the arbitrariness of criterion 2 is proved.

4 th countercheck (13th November, 2002) **Confutation of criterion 5**
Subdivision according to order in finger movement
Figure 8: 4th countercheck version A and **Figure 9: 4th countercheck version B**
When this countercheck was started, 15 principal handshapes and 7 subgroups were selected. The principal handshapes were singled out according to criterion 2 (i.e. moving from the fist to the open hand, followed by the sequential appearance of fingers from 0 to 5), but considering handshapes with joined fingers as subordinate. Within the subgroups criterion 4 (i.e. progressive closure of the fingers but without distinguishing straight and curved finger positions) was followed in contrast with criterion 5. It was therefore proved that, without criterion 5, especially in the "5th finger" subgroup, the attempt to create a sequence is confused since it is difficult to clearly identify "more open" or "more closed" handshapes. Criterion 5 is therefore fundamental. (This countercheck proved to be similar to version 3b in many respects, but it was useful in verifying the importance of criterion 5). In the version following this countercheck, criterion 2 was mainly followed, thereby distinguishing as principal handshapes both the ones with united fingers and the ones with open fingers (e.g. "U" vs. "V") as a movement of maximum opening. Therefore 20 principal handshapes were singled out. Moreover criterion 5 was also taken into consideration. In fact, this countercheck produced subgroups which were very similar to the "3b" version, with a few minor changes. What makes it different from version "3b" are the principal handshape families, created according to the appearance of fingers: "fist" family, "1st finger" family, "2nd finger" family, "3rd finger" family, "4th finger" family, "5th finger" family, which could prove useful for better categorizing and memorizing handshapes. But the negative consequence lies in the additional passages that must be carried out to reach the desired handshape, which could be a further source of confusion. In the present state-of-the-art we have proved that criteria 1, 3 and 5 are fundamental, while criterion 2 is arbitrary.

3. Bibliographical References

Baker, C. & Cokely, D. (1980). *American Sign Language: a teacher's resource text on grammar and culture*. Silver Spring (USA): T.J. Publishers Inc.

Fischer, S.D. & Siple, P. (1990). *Theoretical Issues in Sign Language Research*, Vol. 1. Chicago (USA): The University of Chicago Press.

Kanda, K. (1994). "A computer dictionary of Japanese Sign Language" in Ahlgren, I., Bergman, B. & Brennan, M., *Perspectives on Sign Language Usage*. Durham (GB): The International Sign Linguistics Association.

Kyle, J.G. & Woll, B. (1985). *Sign Language: The study of deaf people and their language*. Cambridge: Cambridge University Press. Lucas, C., Bayley, R. & Valli, C. (2001). *Sociolinguistic Variation in American Sig Language*. Wahington DC (USA): Gallaudet University Press.

Radutzky, E. (1992). *Dizionario bilingue elementare della lingua italiana dei segni*. Roma (I): Edizione Kappa.

Volterra, V. (1987). *La Lingua Italiana dei Segni: la comunicazione visivo-gestuale dei sordi*. Bologna (I): Il Mulino.

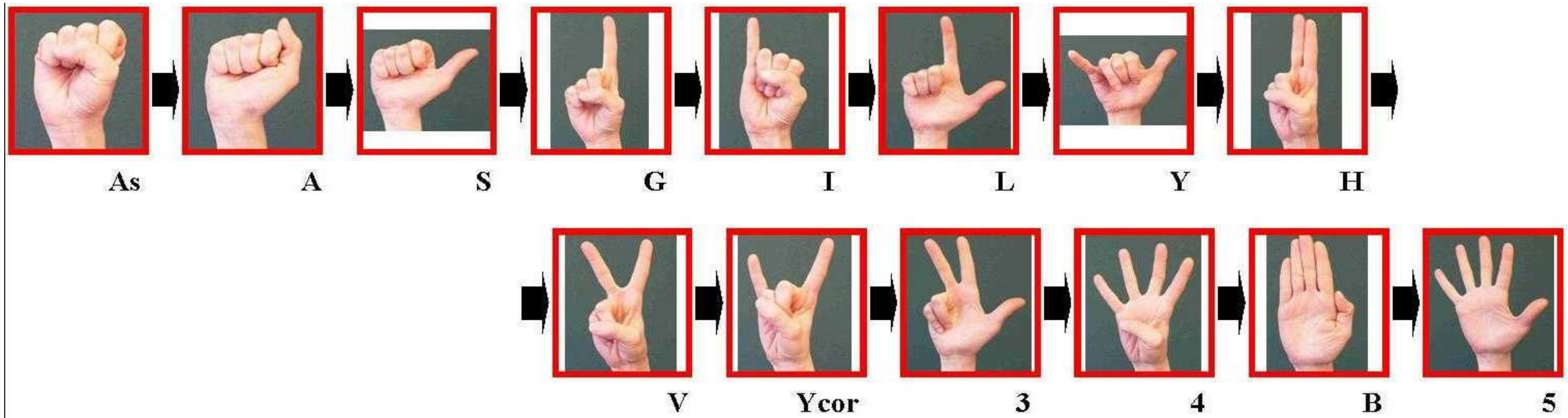


Figure 1: 1st version

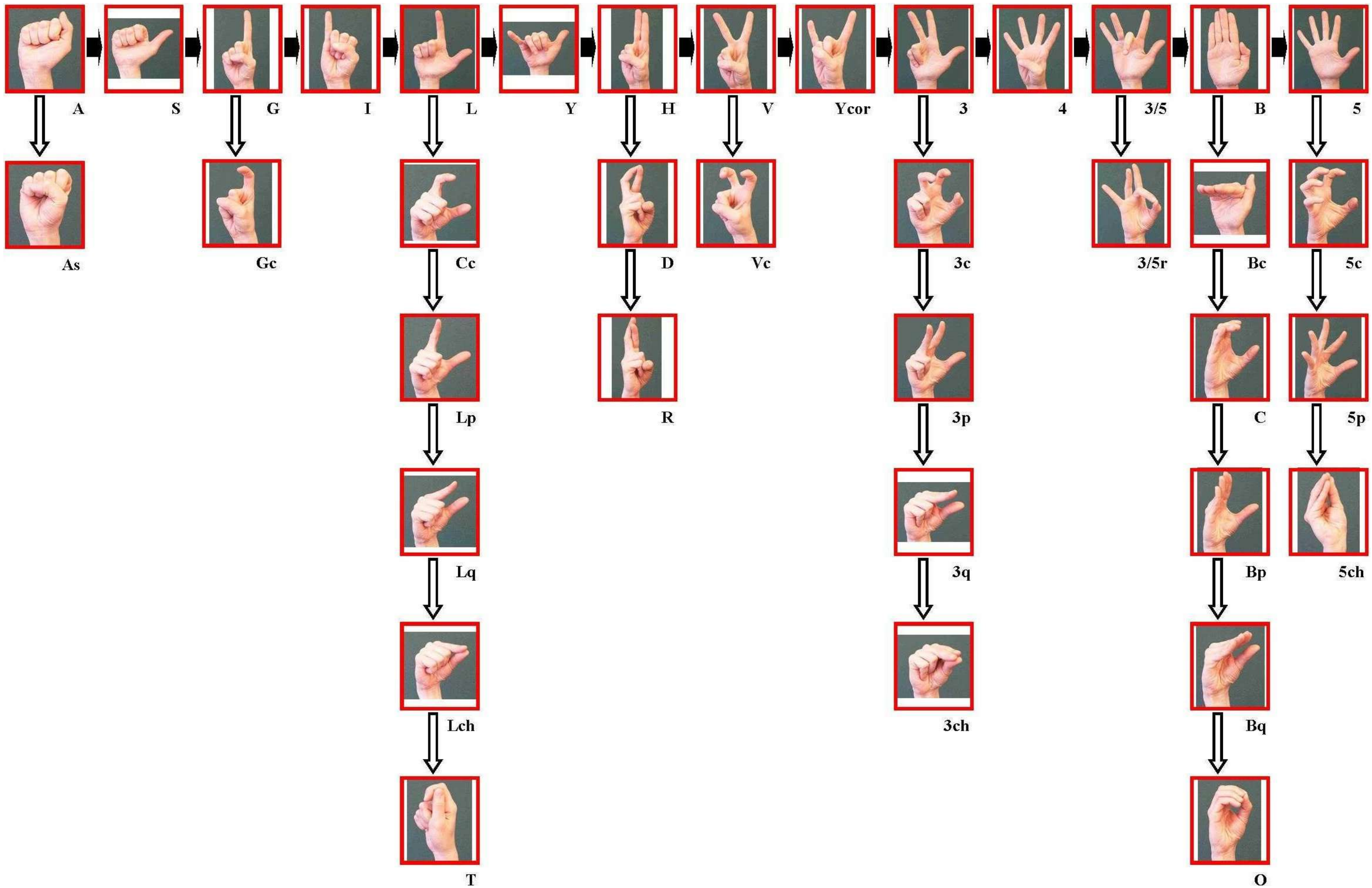


Figure 2: 2nd version

Figure 3: 3rd version A

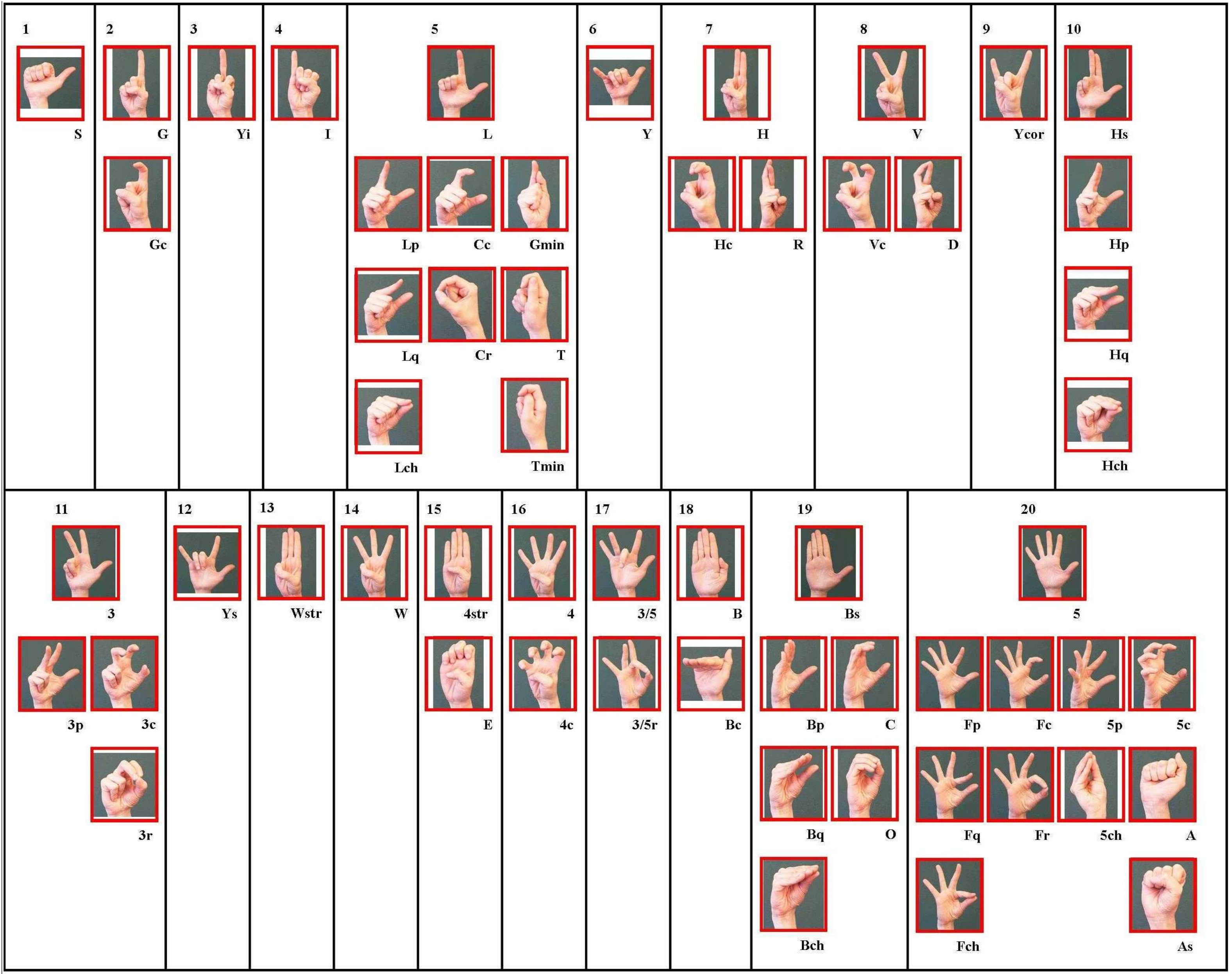


Figure 4: 3rd version B

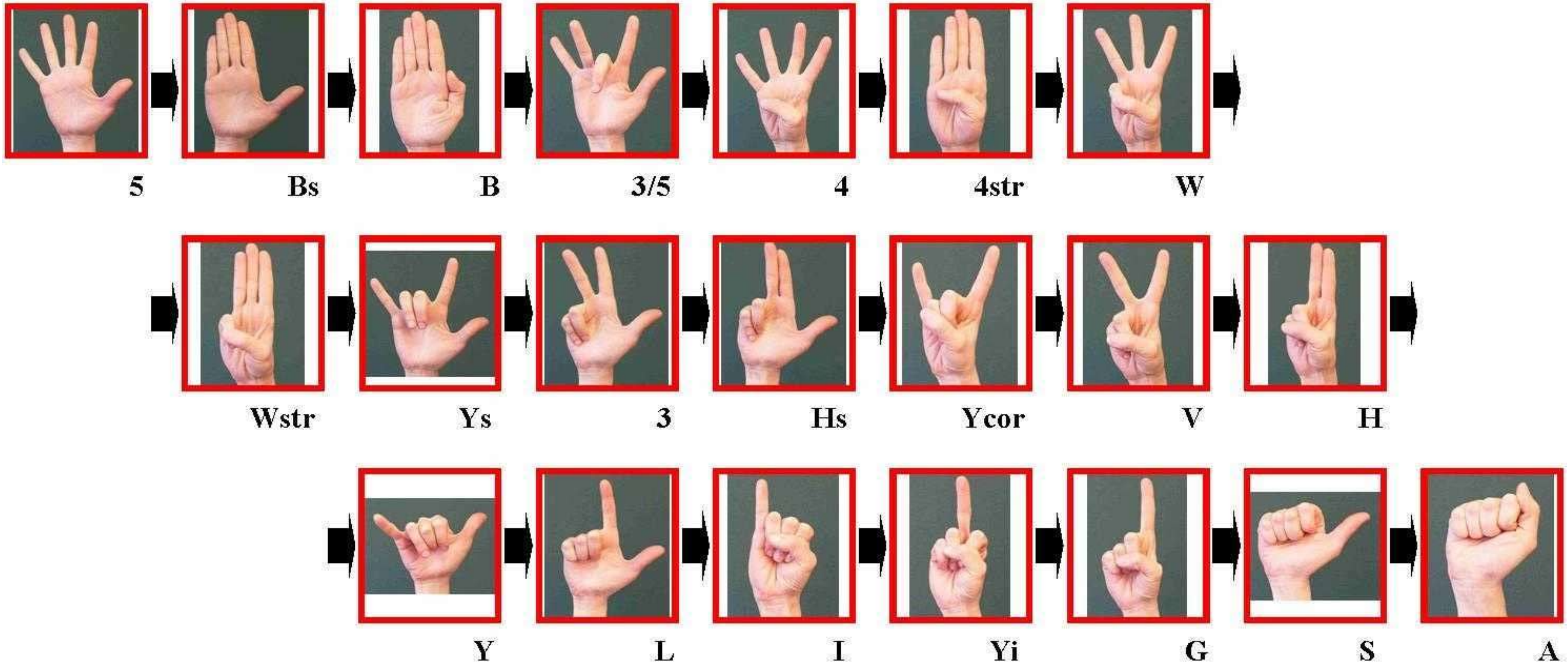


Figure 5: 1st countercheck

Figure 6: 2nd countercheck

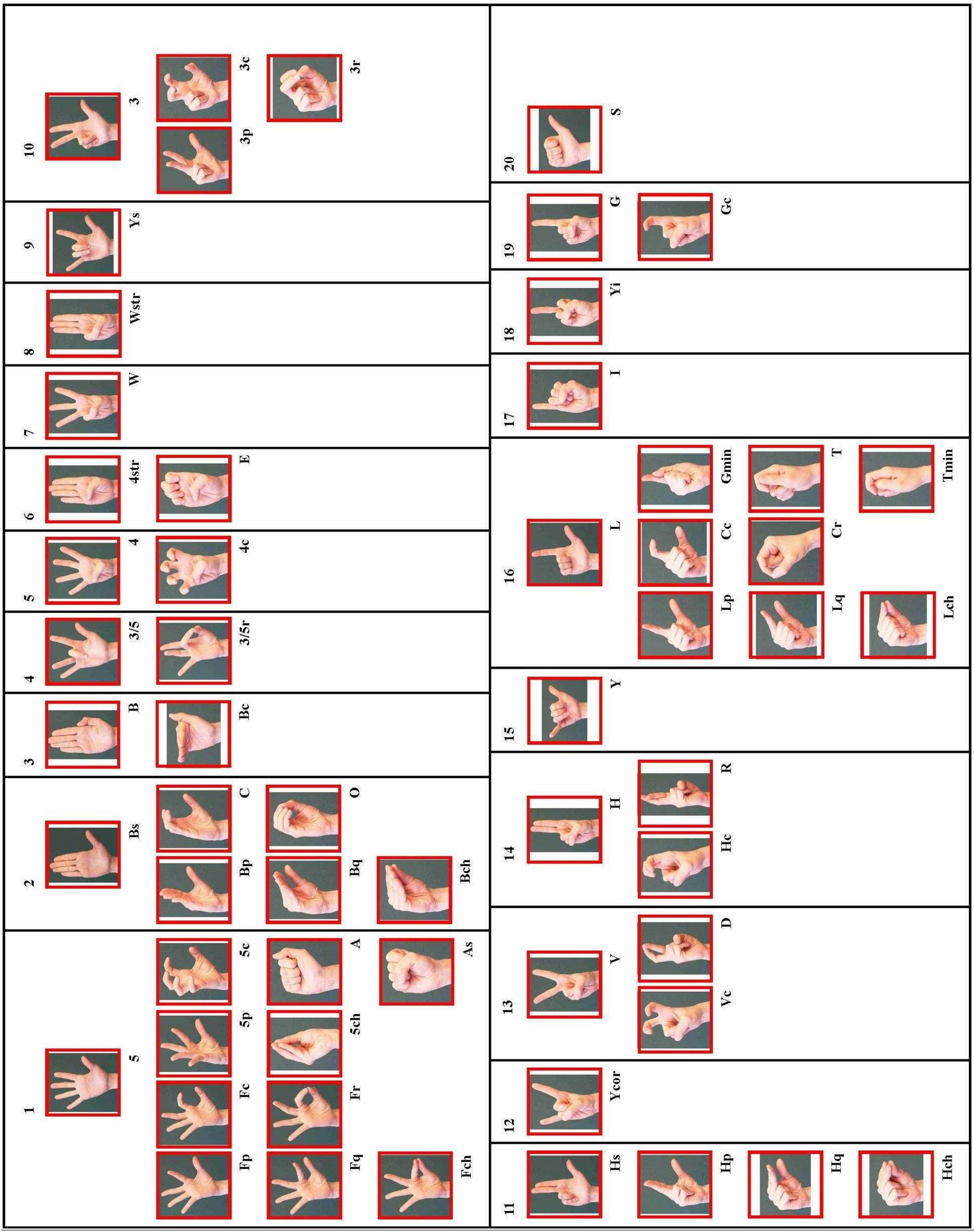


Figure 7: 3rd countercheck

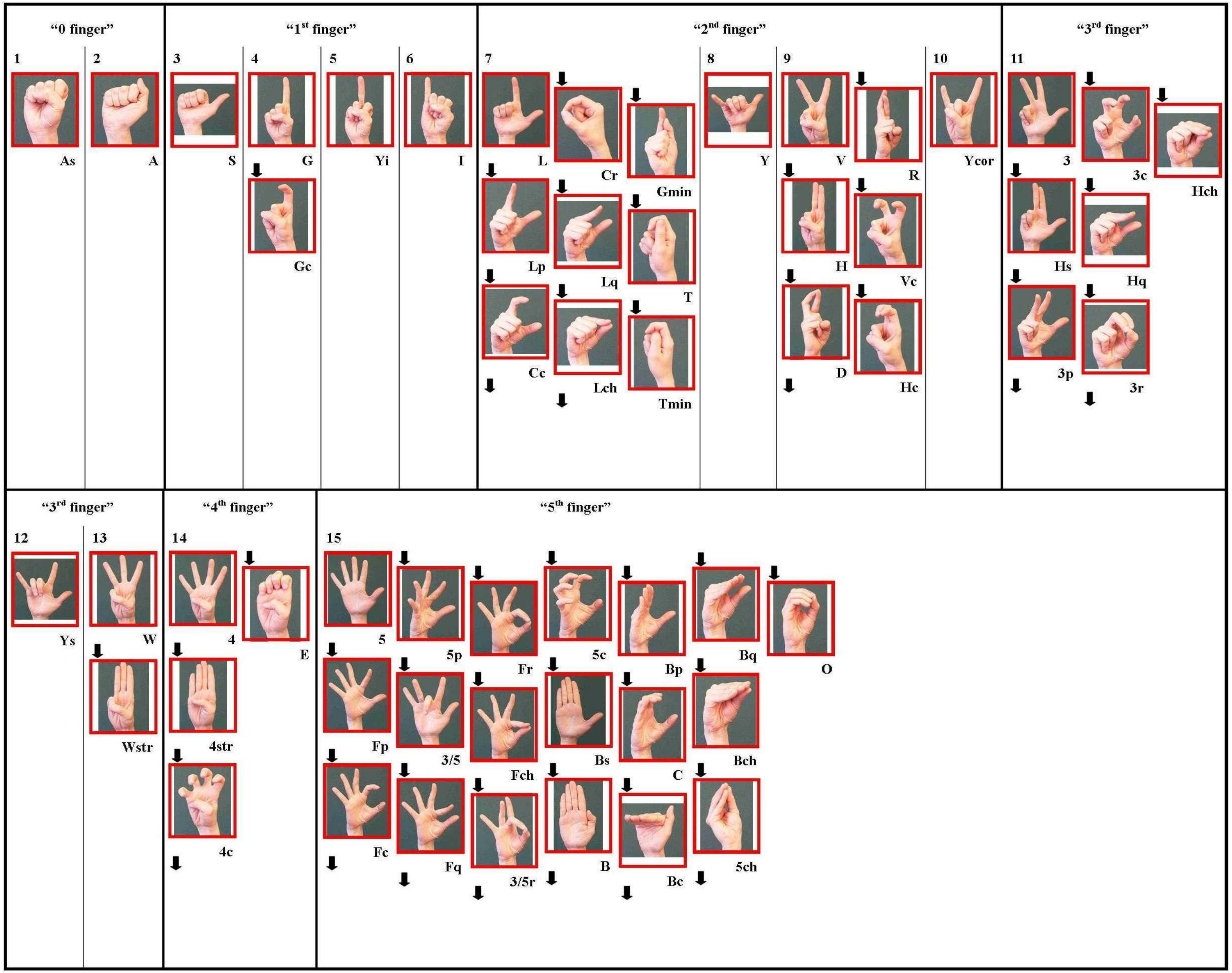


Figure 8: 4th countercheck version A

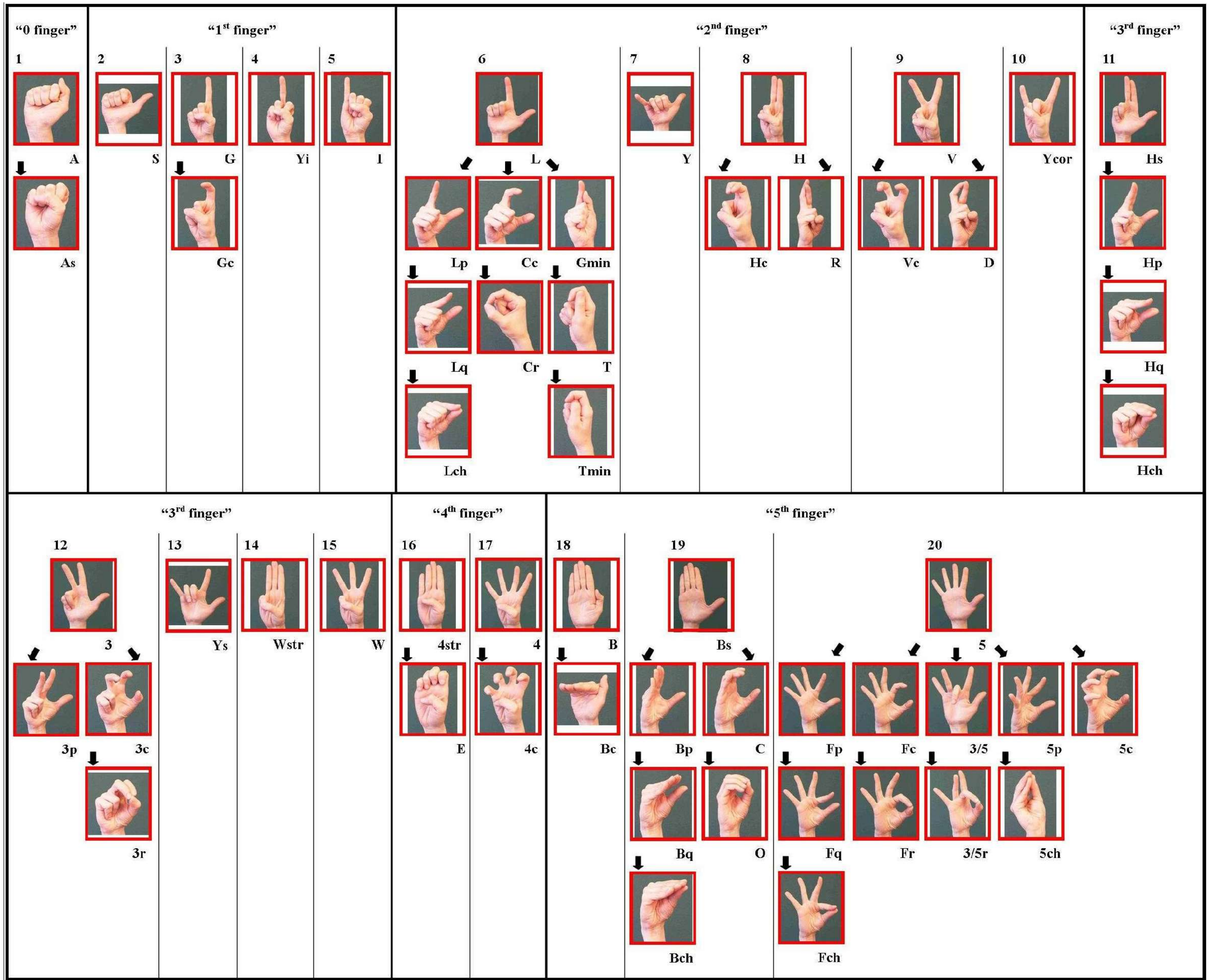


Figure 9: 4th countercheck version B