A SignWriting-Based Approach to Sign Language Processing

Antônio Carlos da Rocha Costa and Graçaliz Pereira Dimuro

Escola de Informática, Universidade Católica de Pelotas
96.010-000 Pelotas, RS, Brazil
{rocha,liz}@atlas.ucpel.tche.br
http://gmc.ucpel.tche.br

Abstract. The SignWriting system is a practical writing system for deaf sign languages, composed of a set of intuitive graphical-schematic symbols and simple rules for combining them to represent signs. The SignWriting Markup Language (SWML) is an XML-based language for encoding sign language texts, written in SignWriting, in an application and computer platform independent way. Thus, sign language texts, written in SignWriting and encoded in SWML, can be entered as input to - and also got as output from - any kind of computer program performing any kind of language and document processing (storage and retrieval, analysis and generation, translation, spell-checking, search, animation, dictionary automation, etc.). This opens the whole area of text-based natural language processing and computational linguistics to the deaf sign languages. The paper explains such approach to sign language processing, giving its elements and general guidelines [1].

Keywords: Sign language processing, writing systems for sign languages, SignWriting, SWML.

1 Introduction

The SignWriting system [2] is a practical writing system for deaf sign languages. It is composed of a set of intuitive grahical-schematic symbols and of simple rules for combining such symbols to represent signs.

Opposed to notation systems like the HamNoSys system [3], SignWriting is not a system specially meant to be used by linguists in their analytical representation of signs and sign phrases, although it can surely be used in such task. It is essentially a system conceived to be used by deaf people in their daily lives, for the same purposes hearing people commonly use written oral languages (taking notes, reading books and newspapers, learning at school, making contracts, etc.)

[1] This work and the SignNet Project (UCPel/PUCRS/ULBRA) are developed with financial support from CNPq/ProTeM and FAPERGS.

[2] http://www.signwriting.org

[3] http://www.sign-lang.uni-hamburg.de/Projects/HamNoSys.html

This places the SignWriting system in a privileged position to be taken as the preferred writing system for sign language and sign document processing systems, as such systems can thus be put into real practical use by common (deaf) people.

Given the graphical-schematic nature of the SignWriting system, an appropriate encoding of its symbols is necessary, in order to allow the computer storage and processing of sign language document files, as well as the use of written sign languages in interactive control components of computer program interfaces.

That is the purpose of SWML (SignWriting Markup Language [4]), an XML-based language that is being developed to allow the computer-platform independent representation of sign language texts written in SignWriting and to allow, thus, the interoperability of SignWriting-based sign language processing systems.

This paper explains our approach to text-based sign language processing, which takes the SignWriting system as its main foundation and SWML as its main computational tool.

Section 2 gives an overview of the SignWriting system. Section 3 firstly reviews XML and its role as a meta-language providing for computer systems interoperability. Then, it briefly explains the current version of SWML, the role SWML can play for future sign language processing systems, and the relation it has to the SignWriter editor (the first and - currently - major computer program used to create and process SignWriting texts and dictionaries [2]).

Section 4 pictures the overall scenario of SignWriting-based sign language processing, as envisioned by the approach proposed here, and hints on software developments being carried on within the SignNet Project [5].

2 The SignWriting System

2.1 Conceptual foundations

Valerie Sutton, the inventor of the SignWriting system, took the stance that, from a practical and intuitive point of view, sign language notation should be visually driven and graphically displayed. Such stance came from her previous experience with the development of a writing system for choreographic movements, the DanceWriting system [4] [6].

Sign language notation was, thus, conceived as just another case of *movement writing*, so that the same principles of DanceWriting could be applied, and the SignWriting system [1] came up as a *visual language* for writing sign languages.

Sure, the system was construed to tackle the various phonetic and phonological aspects of sign languages, as they are usually identified by the mainstream of sign language linguistics [5]: hand configurations, hand and finger movements,

[4] http://swml.ucpel.tche.br
[5] http://sign-net.ucpel.tche.br
[6] http://www.dancewriting.org

locations, face expressions, contacts, segmentation, etc. [1]. That was necessary because the *visual* aspects of sign languages are precisely what is specific to their linguistic features at the phonetic and phonological levels.

However, in its conceptual foundation, the system was kept as a *movement writing* system, and that is exactly what makes it intuitive and usable for common people, not specially trained in linguistics. Also, that is what makes the `SignWriting` system neutral with respect to the alternative linguistic frameworks, and thus compatible with otherwise linguistically incompatible theories.

For instance, the `SignWriting` system is neutral with respect to the various ways to analyze *timing aspects* (sequentiality, simultaneity) in sign language phonology [5], and thus is neutral with respect to the *movement-hold segmentation* versus *single segmentation* debate [6].

In fact, the `SignWriting` seems to be highly compatible with the *visual phonological* approach introduced by Linda Uyechi in [6], which was developed well after `SignWriting` was invented.

2.2 The Graphical Symbols

There are various groups of graphical symbols in the `SignWriting` system, each corresponding to some important (phonetic/phonological) aspect of sign languages.

The figures below (available from the website version of Amy Rosenberg's thesis [12], who collected them from various sources) illustrate the symbols of the `SignWriting` system.

Figure 1 shows the way the system represents basic handshapes. Figure 2 shows the modifications the basic handshape symbols may be submitted to, in order to represent different hand orientations. It also gives examples of American Sign Language signs written in `SignWriting`. Figure 3 is an extract of an ASL text about ASL grammar, written by Karen van Hoek [13]

3 SignWriting Markup Language

3.1 XML and the Interoperability of Computer Systems

The development of the `Internet` furthered the need for the interoperability of on-line systems, and the `eXtensible Markup Language` (`XML`) is the solution proposed by the `World Wide Web Consortium` (`W3C`) to such problem [7].

`XML` is a meta-language allowing the definition of platform- and application-independent languages, dedicated to the storage and processing of information on the Web.

The flexible set of rules incorporated in `XML`, and the wide availability of both free and commercial software (parsers, checkers, validators, etc.) supporting it, as well as the strong commitment to the language by the main computer

[7] `http://www.w3.org/XML`

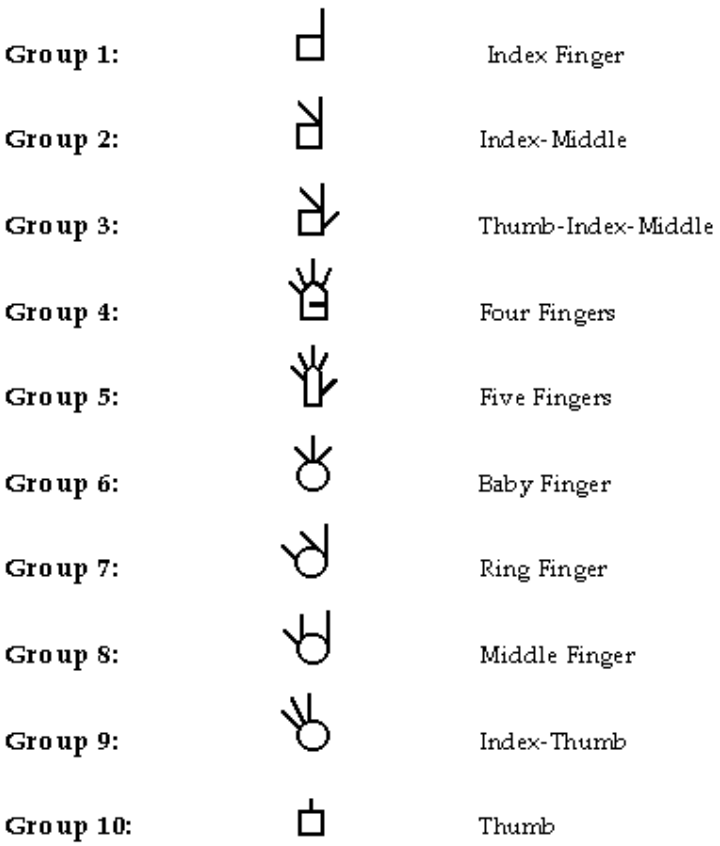

Fig. 1. The ten basic handshapes.

manufacturers and software vendors, turned `XML` into the favorite interoperability tool in every software development initiative concerned with that matter.

As it is easily envisioned that the applicability of `SignWriting` is of a wide spectrum in the Internet (texts, databases, websites, textual conversations, etc.), the need of an `XML`-based format to represent `SignWriting` files can also be easily understood. The `SWML` format explained below attempts to fill such need.

3.2 SWML

The `SignWriting Markup Language` (`SWML`) is an `XML`-based language created to allow the interoperability of SignWriting-based sign language processing systems.

The current version of `SWML` is `version1.0-draft2`, defined by the DTD available at `http://swml.ucpel.tche.br/dtd-version1.0-draft2.htm`.

Its main features are the following:

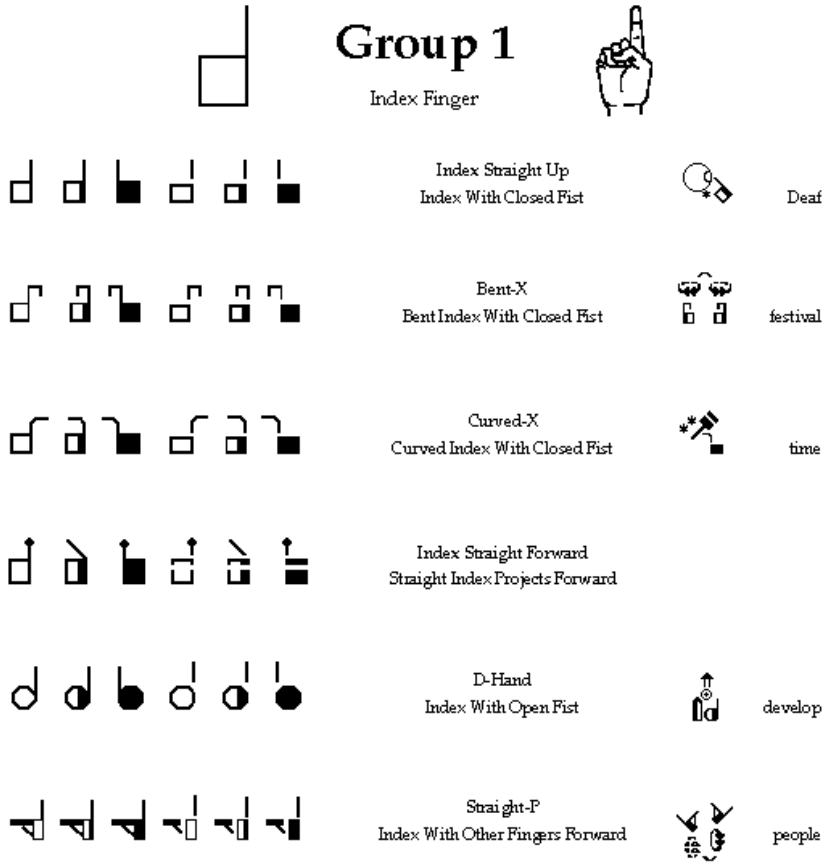

Fig. 2. The various modifications of the Index handshape.

- ◇ `SWML` can represent both `SignWriting` texts and dictionaries, as they are generated by the `SignWriter` program.
- ◇ For every sign in the text or dictionary, there is a `<sign_box>` comprising the set of `<symbol>`s that together represent the sign.
- ◇ For every `<symbol>` in a `<sign_box>`, a `"number"` attribute identifies the `<shape>` of the symbol, and attributes `"x"` and `"y"` identify the coordinates of the symbol within the `<sign_box>`.
- ◇ Besides, for every symbol, a set of attributes (`"variation"`, `"fill"` and `"rotation"`) identify the `<transformation>`s to which the symbol was subjected when included in the sign.
- ◇ The final result is that the various features of any sign can be extracted from its representation in an `SWML`-encoded file, and its image reconstructed, if necessary.

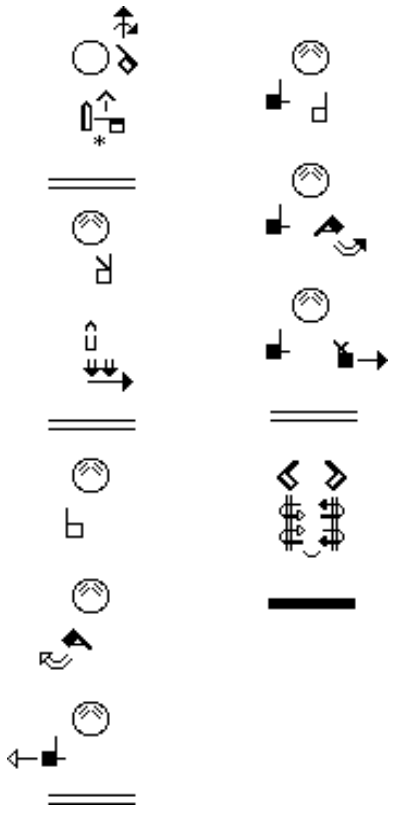

Fig. 3. An extract of the *ASL Grammar Lessons*, by Karen van Hoek, written in ASL.

Figure 4 shows the sign for BRAZIL in the Brazilian Sign Language (LIBRAS). Figure 5 shows the `SWML` file corresponding to that sign, as it is stored in a `SignWriting` text file generated by the `SignWriter` editor. Comments were added afterwards, to ease the reading of the file.

4 SignWriting and Sign Language Processing

4.1 The Proposed Approach

We use the term *sign language processing* to denote the attempt to apply methods and techniques of *natural language processing* and *computational linguistics* to deaf sign languages.

Such methods were originally developed to process oral languages and were, since the beginning, strongly connected to - and even dependent on - methods and techniques of processing oral sentences and discourses presented in *written*

Fig. 4. The sign for BRAZIL in Brazilian Sign Language (LIBRAS).

form. That was a natural start, given (1) the easiness with which oral (Western European) languages could be represented in computer systems, with the Roman alphabet embedded in the ASCII code, and (2) the socially determined dominance of oral languages.

The extension of that work to non-Western European languages posed (and still poses) interesting and defying technical problems, but has not changed the conceptual foundation of the area, because it still targets only oral languages.

The Gesture Workshop series [7,8,9] is one of the forums where an alternative goal for natural language processing has shown up, namely, to consider the problem of processing gestures and sign languages.

That work started by dealing with sign language captured visually, in videos or in real time, which was also a natural start, given the lack of standardized (i.e., universally accepted) written form for sign languages.

Some works presented in those workshops dealt with notations for sign languages (e.g., [10,11]) but the notations were either linguistically oriented (e.g., based on the *Stokoe* system or on `HamNoSys`) or computationally oriented (i.e., modeled after some programming language).

No work, up to now, showed up attempting to process sign language texts as they may be originally produced by native signers that have no special training in linguistics, and no work tackled the problem of common (deaf) user interaction with computer programs using written signs.

Such kind of work, which may well make evident interesting problems concerning the conceptual framework of natural language processing methods and techniques, can only come up with the help of concepts and tools similar in style to the `SignWriting` system and `SWML`.

To set the stage for such kind of work is that we have engaged in the area of sign language processing using the approach explained in the present paper. Very simple computer programs and tools are being developed, as shown below, in order to hint on the conceptual problems that should be tackled in the future, when more sophisticated sign language processing systems may be conceived and proposed.

4.2 Software Development in the SignNet Project

A few simple computer programs and tools are beign developed in the SignNet Project, in order to hint on the conceptual problems that should be tackled by

```
<?xml version="1.0" ?>
<swml version="1.0-d2" symbolset="SSS-1995">
   <generator>
      <name>SignWriter</name>
      <version>4.3</version>
   </generator>
   <sw_text>
      <sw_text_defaults>
         <sign_boxes>
            <unit> pt </unit>
            <height> 60 </height>
         </sign_boxes>
         <text_boxes>
            <box_type> graphic_box </box_type>
            <unit> pt </unit>
            <height> 60 </height>
         </text_boxes>
      </sw_text_defaults>
      <new_line/>
      <sign_box>
         <!-- the B hand -->
         <symbol x="8" y="13">
            <shape number="21" fill="1" variation="1" />
            <transform flop="0" rotation="0" />
         </symbol>
         <symbol x="7" y="25">
         <!-- the movement -->
            <shape number="108" fill="0" variation="1" />
            <transform flop="1" rotation="4" />
         </symbol>
      </sign_box>
   </sw_text>
</swml>
```

Fig. 5. The SWML file for the sign BRAZIL.

more ambitious sign language processing systems, which are to be conceived and constructed in the future.

Some of software developments and algorithm explorations that are being carried on in the Project are the following:

- ◇ New tools do generate and exhibit sign language texts either as stand-alone applications (new `SignWriting` editors) or in web-based environments (web-pages written in sign languages, sign language electronic mail and chats).
- ◇ On-line dictionary systems for sign languages, with information retrieval based on the written forms of the signs.
- ◇ Sign language document repositories, with document search and retrieval based on pieces of written sign language texts.
- ◇ Automated animation of written sign language sentences and dialogs.

5 Conclusion

A `SignWriting`-based approach to sign language processing is possible. Such approach requires a means to guarantee the interoperability of the sign language processing systems based on it. The `SWML` file format is one such means.

From the point of view of the common (deaf) computer user, such approach may be highly practical and useful, since `SignWriting` needs no special linguistic training for its use, requiring only that the user learn how to read and write her sign language in such system.

As the `SignWriting` system was created to be a writing system for daily use, the approach to sign language processing proposed here seems sound with the system's original intention.

Basic computer programs for processing written sign languages are being developed, to add to the sole existing sign language processor, the `SignWriter` program.

As the set of such programs evolve, and users effectively trained in reading and writing sign languages with `SignWriting` progressively produce growing amounts of sign language texts, and also progressively feedback their experiences in interacting with computers using written sign languages, the stock of sign language processing problems will grow, and assessment of the validity of currently available natural language processing methods and techniques will be possible.

References

1. Sutton, V. *Lessons in SignWriting - Textbook and Workbook*. Deaf Action Committee for SignWriting, La Jolla, Ca., 1999. (2nd ed.)
2. Sutton, V. and Gleaves, R. *SignWriter - The world's first sign language processor*. Deaf Action Committee for SignWriting, La Jolla, Ca., 1995.
3. Sutton, V. *Sutton's American Sign Language Dictionary*. Deaf Action Committee for SignWriting, La Jolla, Ca., 1995.

4. Sutton, V. *Sutton Movement Shorthand: a Quick Visual Easy-to-Learn Method of Recording Dance Movement - Book One: The Classical Ballet Key*. The Movement Shorthand Society, Irvine, Ca., 1973.
5. Valli, C. and Lucas, C. *Linguistics of American Sign Language - an Introduction*. Gallaudet Universtity Press, Washington D.C., 1995. (2nd. ed.)
6. Uyechi, L. *The Geometry of Visual Phonology*. CSLI Publications, Stanford, Ca., 1996. (Dissertations in Linguistics Series).
7. Harling, P. and Edwards, A. (eds.) *Progress in Gestural Interaction*. Springer-Verlag, London, 1997. (Proc. Gesture Workshop, GW'96. University of York, March, 1996.)
8. Wachsmuth, I. and Fröhlich, M. (eds.) *Gesture and Sign Language in Human-Computer Interaction*. Springer-Verlag, Berlin, 1998. (LNAI 1371 - Proc. Intl. Gesture Workshop, GW'97. Bielefeld, Germany, September 1997.)
9. Braffort, A. et al. (eds.) *Gesture-Based Communication in Human-Computer Interaction*. Springer-Verlag, Berlin, 1999. (LNAI 1739 - Proc. Intl. Gesture Workshop, GW'99. Gif-sur-Yvette, France, March 1999.)
10. Lebourque, T. and Gibet, S. *A Complete System for the Specification and the Generation of Sign Language Gestures*. In: [9], p. 227-238.
11. Vogler, C. and Metaxas, D. *Toward Scalability in ASL Recognition: Breaking Down Signs int Phonemes*. In: [9], p.211-224.
12. Rosenberg, A. *Writing Signed Languages - In Support of Adopting an ASL Writing System*. Dept. of Linguistics, Univ. of Kansas, 1999. (Master's Degree Thesis, available on-line at `http://www.signwriting.org/forums/research/rese010.html`).
13. Hoek, K. v. *ASL Grammar Lessons*. Written in ASL, with glosses in English. Published in a `SignWriting` file accompanying the `SignWriter` shareware program distribution [2]. Also, reproduced in [12].