Workshop Programme

9:00 – 9:15	**Opening greetings**
	Oral session
9:15 – 9:45	Elena Pizzuto, Paolo Rossini, Tommaso Russo *Representing Signed Languages in Written Form: Questions that Need to be Posed*
9:45 – 10:15	Michael Filhol, Annelies Braffort *A Sequential Approach to Lexical Sign Description*
10:15 – 10:45	Alessio Di Renzo, Luca Lamano, Tommaso Lucioli, Barbara Pennacchi, Luca Ponzo *Italian Sign Language (LIS): Can We Write it and Transcribe it with SignWriting?*
	11:00 – 11:30 Coffee Break
11:30 – 12:00	Patrice Dalle *High Level Models for Sign Language Analysis by a Vision System*
12:00 – 12:20	Morteza Zahedi, Philippe Dreuw, David Rybach, Thomas Deselaers, Jan Bungeroth, Hermann Ney *Continuous Sign Language Recognition – Approaches from Speech Recognition and Available Data Resources*
12:30 – 13:00	Ignazio Infantino, Riccardo Rizzo, Salvatore Gaglio *A Software System for Automatic Signed Italian Recognition*
	13:00 – 14:30 Lunch
	Oral session
14:45 – 15:15	Brigitte Garcia *The Methodological, Linguistic and Semiological Bases for the Elaboration of a Written Form of French Sign Language (LSF)*
15:15 – 15:45	Cynthia Kellett Bidoli *Glossary Compilation of LSP Including a Signed Language: a Corpus-based Approach*
15:45 – 16:15	Gabriel Pereira Lopes *Language Learning by Machines from Plain Raw Text*
	16:30 – 17:00 Coffee Break

17:00 – 19:00 **Poster session**

Rúbia Medianeira Denardi, Paulo Fernando Blauth Menezes, Antônio C. R. Costa
*An Animator of Gestures Applied to Sign Language*s

Eleni Efthimiou, Stavroula-Evita Fotinea, Galini Sapountzaki
Processing Linguistic Data for GSL Structure Representation

Antônio C. R. Costa, Graçaliz P. Dimuro, Benjamin C. Bedregal
Recognizing Hand Gestures Using a Fuzzy Rule-Based Method and Representing them with HamNoSys

Guylhem Aznar, Patrice Dalle
Analysis of the Different Methods to Encode SignWriting in Unicode

Maria Mertzani
Sign Language Learning through Asynchronous Computer Mediated Communication

Thomas Hanke
Towards a Corpus-based Approach to Sign Language Dictionaries

Sandy Pleissner
Translation of Natural Speech into Sign Language Based on Semantic Relations

Annelies Braffort
Articulatory Analysis of the Manual Parameters of the French Sign Language Conventional Signs

Onno Crasborn, Han Sloetjes, Eric Auer, Peter Wittenburg
Combining Video and Numeric Data in the Analysis of Sign Languages within the ELAN Annotation Software

Maria Mertzani, Clark Denmark, Linda Day
Forming Sign Language Learning Environments in Cyberspace

Emilio Insolera, Maria Giuseppina Militano, Elena Radutzky, Alessandra Rossini
Pilot Learning Strategies in Step with New Technologies: LIS and Italian in a Bilingual Multimedia Context 'Tell Me a Dictionary

Emanuela Cameracanna, Maria Luisa Franchi
Metodo Vista – Teaching Sign Language in Italy

Workshop Organiser

Chiara Vettori

Institute for Specialised Communication and Multilingualism

EURAC research, Italy

Workshop Programme Committee

Chiara Vettori

EURAC research, Italy

Antônio Carlos da Rocha Costa

Escola de Informática - UCPel, Brazil

Eleni Efthimiou

Institute for Language and Speech Processing – ILSP, Greece

Elena Pizzuto

Istituto di Scienze e Tecnologie della Cognizione - ISTC, Italy

Elena Radutzky

Mason Perkins Deafness Fund, Italy

Marco Consolati

ALBA s.c.s., Italy

Onno Crasborn

Dept. of Linguistics - Radboud University Nijmegen, Netherland

Patrice Dalle

IRIT - Université Paul Sabatier, France

Ronice Müller de Quadros

Centro de Ciências da Educação, UFSC – Universidade Federal de Santa Catarina, Brazil

Tania Di Mascio

Department of Electric Engineering, University of L'Aquila, Italy

Table of Contents

Author Index

Representing signed languages in written form: questions that need to be posed

Elena Pizzuto[1], Paolo Rossini[2], Tommaso Russo[3]

[1]Istituto di Scienze e Tecnologie della Cognizione - CNR (ISTC-CNR), [2]Istituto Statale Sordi di Roma (ISSR)
Via Nomentana, 56 - 00161 Roma, Italia
[3] Dipartimento di Filosofia, Università della Calabria, Cosenza, Italia
E-mails: elena.pizzuto@istc.cnr.it, paolo.rossini@istc.cnr.it, t.russo@mclink.it

Abstract

In this paper we discuss some of the major issues linked to the unwritten status of signed languages and to the inadequacy of the notation and transcription tools that are most widely used. Drawing on previous and ongoing research, we propose that the development of a written form appears to be necessary for defining more appropriate representational tools for research purposes.

1. What we need to consider on the road towards electronic representations of Signed Languages

The purpose of this paper is to point out, and briefly discuss, some key questions which we believe must be clarified, or at least explicitly formulated, *prior* to focusing on the main issue this workshop aims tro address, i.e. how to represent Signed Languages (SL) electronically[1]. Our observations are based on both previous and ongoing research conducted primarily on Italian Sign Language (LIS) within our group (Fabbretti & Pizzuto, 2000; Pennacchi et al, 2001; Pietrandrea, 2000; Pizzuto & Pietrandrea, 2001; Pizzuto et al., 2000a; Russo, 2000), and on recent work undertaken within a comparative crosslinguistic project on LIS and French Sign Language (LSF) (Pizzuto & Cuxac, 2004; Garcia & Dalle, 2005) . However, the poblems discussed below are not language-specific, and constitute a major obstacle for developing appropriate cross-linguistic investigations of sign languages (Pizzuto et al., 2000b; Pizzuto & Cuxac, 2004).

Our considerations are grounded upon a more general theoretical-methodological and socio-cultural framework, sketchily outlined in this section, and should be related to those formulated, in the same line of thinking, by Garcia and by Di Renzo et al (in this volume).

We think that in order to devise more appropriate means for representing SL it is necessary, in the first place, that hearing and Deaf researchers working on SL adopt a broad semiotic and socio-cultural perspective for defining and describing SL, and the communities of Deaf signers, much more clearly than it has been done so far. Most past and current research provides an oversimplified view of SL and Deaf communities. For example, though of course everybody recognizes that 'native signers' constitute an extremely small proportion (5% or less) of all signers, most descriptions of institutionalized, national SL are based on data drawn from this very small sample of SL users.

By comparison, very little attention is devoted to characterize accurately the extremely complex linguistic, sociolinguistic and sociocultural variation that is found among SL users, due to the fact that the remaining, vast majority of signers (95% or more) are not native, and acquire their national SL at very different ages, and in very different conditions. Non-institutionalized SL, such as 'home-sign systems', and Primary or Emergent SL used by Deaf individuals who grow out of contact with other Deaf persons, or who develop SL in micro-communities, also need to be taken into account (Fusellier-Souza, 2004). This rich variation needs to be explored and assessed very carefully for defining the 'object' we wish to represent: do we want to represent only, or primarily, the variety used by an extremely small sample of users, or the many varieties that we know exist?

Second, it seems to us equally necessary that Deaf signers at large, and Deaf researchers in particular, become much more significantly involved in the scientific and cultural issues at stake. The active participation of Deaf signers in research teams working on SL (an obvious 'must' for any serious team) is not sufficient, per se, to ensure that the issues we face are addressed appropriately. What is needed is that Deaf researchers contribute to articulate in new and meaningful directions a much needed discussion of major issues that are crucial for developing written representations of SL. Within the limits of the present paper, we briefly discuss only two such issues: (a) the distinctions between languages with and without a written tradition; (b) the differences, and relations, between writing and transcribing.

1.1. Written vs. unwritten status, writing vs. transcribing

It is unquestionable that, until now, none of the SL used in the world has autonomously developed a written form. Different writing codes have been proposed, since Bebian's (1825) early attempts to "normalize" LSF and, in the last decade especially, a writing system such as Sign Writing (SW), designed by Valerie Sutton (1999), seems to gain a growing consent in several community and educational centres around the world. However, at present we certainly cannot say that wether these "signs of writing" will lead to a real evolution of written SL.

[1] For the sake of clarity, it may be useful to specify that signed language data are already electronically "re-presentable", in a trivial sense of the term, since they can be stored on, reproduced from, and (partially) searched/retrieved, via/on different electronic devices and supports: videotapes, CDs, DVD. The problem we explore here is how to represent signed languages via appropriate written and/or graphic codes that mirror, and allow us to reproduce SL form-meaning patterns *indirectly*, via conventional codes functionally comparable to those available for storing, analyzing and dexcribing spoken/written data.

SL must thus be assimilated, under many respects[2], to spoken languages with an oral-only tradition, learned and transmitted in face-to-face interaction. As recalled by Di Renzo et al (this volume) this feature is not unusual because it is proper of the vast majority of human spoken languages, but the crucial differences between "oral" and "written" languages and communication need to be taken fully into account (Halliday, 1985; Ong, 1982, among others).

The "oral" or "face-to-face" status of SL is relevant with respect to the question of variability of SL lexical and grammatical structures. Primarily oral languages lack one of the main drive toward a "common standard" i.e.written texts. Written texts and written grammars afford socially approved models of "well formed" language structures and thus greatly contribute to language standardization processes. In developing written representation of SL, the issue of variability is clearly at stake. Which forms and structures, which of the different lexical and grammatical variants are to be codified in a written form? It is obvious that this question can be answered only by Deaf signers, and the related "norms" require extensive involvement of Deaf communities (see Garcia, this volume).

For all the language communities that use it, a *writing system* is a socially shared code employed for the transmission of texts, overcoming time and space limitations. Different kind of "communicative needs," can be at the origin of a writing system (e.g. the possibility of fixing and trasmitting a shared body of laws, the transmission of literary texts, the elaboration of written dictionaries and grammars for educational purposes). Writing system are thus created in order to respond to communicative, artistic and educational needs and are designed to do that.

Writing systems undoubtedly provide an analysis of language structures which must be sufficient to achieve particular ends, and which is not an exhaustive analysis of the language structures. The linguistic structures which are codified in a writing system are the ones necessary to vehiculate the meanings which are communicated in particular settings and for particular purposes or usage. Thus different societies and cultures (e.g. the chinese written culture vs. the western tradition) choose different aspects of a language in order to better achieve these ends. This always occurs through a social process of elaboration, diffusion and institutionalization of the writing system.

On the other side *transcription systems* are tools that are useful for researchers in order to analyse linguistic structures, and are developed in order to represent the linguistic features that are studied by the researcher. Most importantly, transcription systems do not influence language usage and language varieties, while writing systems do. Indeed, as noted, writing systems contribute to the standardization processes and thus influence linguistic norms, provinding structures that are to be conceived as a model of a socially approved, "well formed" way of using a language.

The history of writing systems demonstrates that in general the possibility of transcribing texts is always subordinated to the emergence of a writing system. Writing systems involve a particular form of metalinguistic awareness by providing a segmentation of the linguistic structures. This kind of metalinguistic awareness becomes an integral part of language users' linguistic competence, and thus transforms language usage. Transcription systems are facilitated by the emergence of a writing system because writing provides a representation of linguistic competence that is socially shared and commonly agreed upon.

In this frame, and limiting our attention to our "literate" societies, Deaf researchers and signers at large are in a peculiar, culturally disadvantaged position that needs to be highlighted. As remarked in this volume by Di Renzo et al, Deaf signers live in a diglossic environment, in which their unwritten face-to-face SL must co-exist with the dominant spoken and written language used by, and in interaction with, the surrounding hearing community. However, due to well known difficulties engendered by deafness, most Deaf signers, including several highly skilled and qualified Deaf researchers, do not develop appropriate literacy skills in the dominant written language (see also Garcia, this volume).

We cannot underestimate the difficulty these researchers encounter when they try to ground the distinction between writing and transcribing on their own language experience. Their SL is not written, and this fact by itself renders very problematic drawing the distinction between writing and transcribing *with respect to their own SL*. For the same reason, in the absence of a written form, it is certainly not easy, for Deaf researchers, to evaluate the appropriateness of the various "notation" and "transcription" tools that have been proposed for their SL (see section 2 below). One could argue that Deaf researchers can still draw the distinction between writing and transcribing resorting to their knowledge of the dominant spoken/written language which they also use. We may grant that this can be done. Yet the very little discussion that there has been thus far over these themes, along with unfortunate practices that, for lack of better tools, continue to be used in SL research (see section 2), indicate that much remains to be clarified. In any event, the problems pointed out above with respect to SL remain unresolved, and need to be faced.

From a more general perspective, we believe that the complexity of the issues to be faced demands a significant cultural effort, to be carried out jointly by the hearing and Deaf communities involved: we need to create appropriate conditions that allow Deaf signers, and especially Deaf researchers, to have extensive access to, and elaborate, relevant information and theorizing on the problems linked to the representation of SL. We can no longer ignore, nor underestimate the "language barriers" that severely limit Deaf signers's access to much relevant information on their own community.

2. Questions that need to be posed

We turn now to consider specific questions we believe need to be posed with respect to: (a) comparing SL vs. vocal language (VL) corpora; (b) the limits of Stokoe-based notations for representing SL; (c) the misuse of so-

[2] Note that, unlike oral-only languages used in a specific geographical area, SL are characterized by a peculiar diglossic situation, within "literate" societies (see later in this paper), and lack geographic unit within each national community (e.g. there is no "LIS-land" or "LSF-land" comparable to the "x-land" of a geographically delimited oral-only language).

called “glosses” in SL research; (d) the issue of writing vs. transcribing.

2.1. SL vs. VL corpora

There is one implication of the unwritten status of SL that is rarely, if at all, discussed in current research. Given their unwritten status, SL should be naturally assimilated to, and studied as, languages with an “oral-only” tradition (but see footnote 2). At the very least, as remarked in different studies (Fabbretti & Pizzuto, 2000; Pietrandrea, 2000; Pizzuto & Pietrandrea, 2001; Pizzuto & al, 2000a; Russo, 2000; 2004; 2005), an appropriate investigation of SL would benefit from taking seriously into account the frame of reference developed for spoken/oral (as distinguished from written) forms of language (e.g. Biber et al. 1999; Halliday 1985; Ong 1982).

In this frame, our first question is the following.

Q1: How many studies do we know that have taken fully into account the primarily “face-to-face” or “oral” (in the broad sense of the term) status of SL? How many appropriate crosslinguistic and crossmodal comparisons have been performed between SL and VL corpora, most notably *text corpora drawn from actual usage*?

From the answers we can provide to these questions we can measure the gaps in our current knowledge of SL as unwritten languages. The representational issue is obviously crucial.

2.2. The limits of Stokoe-based notations

In spite of the remarkable progresses made since the modern study of SL began with Stokoe’s (1960) seminal work on American Sign Language (ASL), we still do not have efficient, widely standardized notation/transcription tools for representing SL (Bergman et al, 2001). As argued in different studies produced within our group (e.g., Pizzuto & Pietrandrea, 2001; Pizzuto & al, to appear; Russo, 2000), this lack of appropriate tools can be most readily appreciated when we try to represent *signed texts*, or even very short sequences of signs in units characterized (often without clear definitions) as “sentences”, “clauses”, “utterances”.

The main issue to be faced, in our view, is a somewhat paradoxical theoretical and methodological problem. The kind of notation originally proposed by Stokoe has been subsequently employed in the investigation of many (almost all?) SL, with more or less extensive expansions and modifications, and/or significant implementations for the computational representation of SL as for example in the HamNoSys (Prillwitz & al, 1989 -- see also the collection of papers in Bergman et al, 2001 and Streiter & Rocha Costa, 2004 for overviews of current notation tools). However, Stokoe-based notations can be succesfully employed primarily for notating single, decontextualized signs, as in the citation forms listed in SL dictionaries.

But this notation cannot be used for segmenting and transcribing individual signs and signs’ sequences occurring in the actual flow of signed conversation, with all the morphological modifications noted in discourse. The limits of this notation are also evidenced by the fact that, to our knowledge at least, there are no *monolingual* dictionaries or reference grammars that rely on this notation as the primary and only means for representing the signs they describe. The “representation-by-notation” given in such reference tools is not autonomous, but it is always substantially integrated with text descriptions in a specific written language (e.g. English, Italian, Spanish), and graphic, pictured or filmed illustrations of the signs described. These descriptions are in no way comparable to those we find in dictionaries and reference grammars for spoken languages.

Q2 can thus be formulated as follows: are we sure that our analyses of the linguistically relevant manual and nonmanual elements that compose the signs, and allow their organization and segmentation in discourse, are appropriate? Or isn’t rather the case, as suggested in related work (see Pizzuto & Pietrandrea, 2001; Russo, 2005, among others), that the difficulties we find in using Stokoe-based notations for transcribing signed texts reveal a need to revise our current analyses of SL structure much more profoundly and extensively than it is commonly assumed?

The latter view appears more plausible in the light of the following considerations. First, it seems highly peculiar that a notation tool assumed to be adequate for representing isolated, decontextualized “lexical units” cannot be used for representing the same units when they occur in context. This fact in itself should generate “suspicion”, since it seems to have no parallel in spoken language research. For example, if we were to use the IPA notation for representing decontextualized lexical items of a VL that has never been previously described, it seems unlikely that we would be unable to write down the same items when they occur in sequences of spoken discourse.

There is another peculiar phenomenon that can be noted in much lexicographic work, and which may be linked in part to the use of Stokoe-based notation (but also to the unfortunate practice of “glossing” and to the difficulties of constructing dictionaries from corpora of actual usage -- see Brennan, 2001; Russo, 2005; and below). The signs that are included in SL dictionaries are for the most so-called standard signs (though this definition is far from being clear and somewhat circular, since the very inclusion of a sign in a dictionary is one of the element for classifying it as “standard”). The vast majority of such signs turn out to be units which can be easily translated via single words of the contact/dominant language (e.g. by common words such as “bed, sleep, child, table, glass, see, man, woman” etc.)

Typically missing from dictionaries are complex units that are commonly characterized as part of the “productive lexicon” and encode equally complex meanings for which it is often difficult to find single-word translations. These units include manual and nonmanual components, and have been described with different terms, including “classifiers” “classifier-based” or “polymorphemic” predicates, “polycomponential signs” “productive morphemes” (see among others Emmorey, 2003, for a recent overview, and Di Renzo et al, this volume, for some illustrative examples). In recent research we have found more fruitful to characterize these complex units as Highly Iconic Structures (hereafter HIS), adopting, and extending to LIS, the theoretical-methodological framework proposed by Cuxac (2000) for LSF.

Signers’ intuitions and empirical evidence from analyses of fairly large corpora suggest that HIS are a very relevant feature of signed discourse. In research on LSF Sallandre (2003) has found that in some kinds of narrative texts HIS can constitute as much as 70% of the sign units produced. Disregarding terminological differences, work conducted by Brennan (2001) on British Sign Language (BSL), and by Russo (2000; 2004) on LIS provides very similar indications. In addition, both Russo's (2000; 2004) and Sallandre's (2003) studies provide, from different perspectives, important evidence on the large intra-subject variability that characterizes the use of HIS according to different discourse genres and registers.

It is extremely difficult, if at all possible, to capture HIS via a Stokoe-based notation, especially if one wishes to describe accurately the nonmanual components of these complex units. This is an additional indication that a Stokoe-based analysis, and the related notation tools, are not adequate for our descriptions and representations of SL lexical and morphological structure. Since HIS appear to constitute such a relevant dimension of SL structure, these limitations and inadequacies can no longer be overlooked or underestimated.

2.3. The misuse of so-called "glosses"

Our third question concerns the unfortunate yet widespread practice, in SL research, of resorting to so-called "glosses" for parsing and "writing down" what are considered the "basic meanings" of signs identified in signed utterances and texts. The words used for this purpose are in CAPITAL letters by convention (e.g. "EAT" for a sign meaning "eat").

The term "gloss" is actually a misnomer for the labeling operation that is performed in SL research. In fact, glosses as appropriately used in the annotation of spoken/written language data are always an *ancillary* device that *does not replace, but accompanies*, in a reference language known to the author and the reader of a given study, *an independent representation of the language data object of inquiry*. The example in (1) below, taken from Pulleyblank's (1987: 988) description of Yoruba (a Nigerian language) illustrates this point.

(1)	ó	gbé	e	wá
	he/she	carry	it	come
	'He/she brought it'			

The first line in (1) provides an independent, orthographic representation of the constituent units parsed in the Yoruba utterance described. The second line provides, in a one-to-one correspondence, English glosses for the elements represented on the first line, while the third line provides the English translation. This is a plausible and useful use of glosses, as ancillary notation tools that help the reader to understand (via labels in a familiar language) the lexical and morphological patterns of the Yaruba sequence. But the constituent elements of the original sequence are and must be represented independently, otherwise we simply would have no idea of the form-meaning patterns of the language investigated.

Quite differently, in SL research glosses are used as the *primary* and indeed *only* means for representing signs in a written form. For example, a 'glossified' rendition of a LIS sequence with a meaning comparable to that in (1) could be:

(2)	INDEX-a	INDEX-b	BRING

A text (in English, Italian etc.) would then accompany the representation in (2) describing, for example, where the "INDEX" signs were directed and located in space, whether the verb labelled as "BRING" was/was not dislocated in space, and the like. The point is that, in the kind of "representation" provided in (2), the reader has no way to reconstruct the LIS forms that were produced. *There is no independent representation of the signs*, hence nothing is being "glossed". What we have is just "word-labels" for the meanings we assigned to forms that plainly are not "there". By the same token, any so-called transcription of SL data via labels of this sort cannot be defined a "transcription" in any appropriate sense of the term.

The use of word-labels has one other major detrimental effect which has been described quite extensively, and we will only mention it here: these labels can grossly misrepresent the structure of both individual signs and signed discourse (Jouison, 1995; Pizzuto & Pietrandrea, 2001). For example, in research on LIS (Pennacchi et al, 2001; Pizzuto & al, to appear) we have shown how word-labels can lead to inappropriate parsing of utterances within a structured sequence of signs. In earlier work, using a fairly detailed notation based on word-labels, we had analyzed a given sequence of signs as consisting of a single utterance composed of five manual sign units. In subsequent work, the same sequence was transcribed using SW symbols. The SW-based representation of the manual and nonmanual components provided markedly different results, leading to identify three utterances, rather than one. It is important to note that both analyses were perfomed with the substantial help of a highly competent native LIS signer (Rossini, co-author of the present paper). The different results obtained in the two analyses appeared thus to be significantly influenced by the representation tool employed.

The question we want to pose on the ground, then, is the following.

Q3: If we all agree that segmenting and labeling the signs occurring in signed texts via word labels is very inadequate, and even dangerous, why do we continue to do it? Even granting that there are, in fact, practical reasons why this unfortunate practice continues, why is the problem still so widely underestimated in much current research on SL?

2.4. Can writing be bypassed?

The last question we would like to formulate is apparently very simple, but dense of implications for much current research on SL.

Q4: Are we sure that, in our attempts to develop appropriate notation/transcription tools for SL, we can "bypass" the development of some form of writing, before proceeding any further?

As researchers who have been actively involved in the study of SL for rather extensive periods of time we have always been extremely surprised by the lack of interest, when no overt opposition, that most researchers in the field appear to manifest towards the issue of "writing SL". We must admit that we find difficult to understand the reasons of this state of affairs. As noted earlier with respect to spoken languages, the possibility of transcribing texts is rather naturally subordinated to the existence and use of writing systems. It is hard to imagine why this should be different for SL. We think that, as researchers, we should reflect on the question formulated above, and try to address the issue it raises, motivating on theoretical and/or empirical grounds the reasons that may lead us to provide an affirmative or negative answer.

3. Searching keys lost in the dark

Anyone who has confronted him/herself with the task of analyzing and describing meaningful linguistic patterns and structures in SL corpora drawn from actual signed discourse knows from direct experience, as we do, how difficult it is to perform this task with the "transcription" and notation tools currently available, and most widely used. The intellectual uneasiness one experiences is no less relevant, especially with respect to the use of word-labels for "pretending" to represent signs connected in discourse, especially if we try to compare what we do with (and on) SL to what one would normally do, or not do, in VL research.

We believe that no field linguist would try to uncover and describe the lexical and grammatical structure of a VL "X" that has never been explored using, as a major "notation" tool for "fixing down on paper" the patterns of "X", the words of his/her own spoken/written language (e.g. English). But this is exactly what happens, in SL research, everytime we use word-labels to parse signed discourse, and pretend to represent form-meaning patterns via words of a spoken/written language. Since the *signed forms* are in no meaningfull sense represented, no form-meaning patterns are described. In addition, this "representation" can seriously distort and prejudice our analyses. Yet the unfortunate practice of word-labels for signs continues, questioned by many, but apparently not questioned enough to be abandoned. Different, but no less relevant criticisms can be raised with respect to Stokoe-based notations, as we tried to illustrate above.

To use a deliberately provocative metaphor, it seems to us that this way of proceeding could be likened to that of the character of a well-known story who, in a dark night, lost his keys and was searching them under a street-lamp light. Questioned by a passing-by policeman whether he had lost his keys right there, under that street-lamp, he replied: "actually not". Further questioned as to why, then, he was looking under the lamp, the man replied "at least there is some light here".

We believe that the work we have begun on "writing and transcribing LIS signs with the glyphs of SW", reported in this volume by Di Renzo et al, provides indications for at least beginning to face the problems discussed in this paper, while avoiding the fallacy of searching in wrong (though perhaps more familiar) places the keys we need to "unlock" these problems.

We wish here to comment briefly on some aspects of this work that we have found particularly promising. As we observed how signers composed their written texts, and how these were read by other signers, it seemed evident that the written forms produced mirrored the signers' internal competence, and allowed them to express their "LIS-grounded thoughts" directly and effectively, in a way they had never experienced before (e.g. with respect to how signs were parsed, how relevant manual and nonmanual components were selected).

We found of special interest the fact that the written texts included not just so-called standard signs, but also a fair amount of HIS, and that the the SW glyphs could easily represent signs organized in a multilinear fashion, mirroring the coarticulation in space of distinct signs in a single temporal unit that is found in signed discourse.

On the other hand, the use of SW for transcribing signed texts, and the comparisons that could then be done with written texts, allowed signers to quite literally "see" key structural differences between written and face-to-face texts that would have never emerged without a written representation of *both* kind of texts. The distinction between what one "knows" (when producing a written text), and what one actually "does" (when producing a signed text) thus became much clearer, because it could be grounded on a written representation of form-meaning patterns, in the signers' own native language, in different modalities of language expression. This posed the basis for a much deeper, theoretically and empirically motivated understanding of the crucial distinctions between writing and transcribing (see Di Renzo et al, this volume). The discussions we had give us good reasons to think that the insights that were achieved could never have been gained without the help of a written code.

All of this indicates that a system such as SW has the potential for encoding structures and morphosyntactic organizational patterns that are highly specific of SL, and that emerge not only in their face-to-face form but also, and most interestingly for us, in their written form. It also suggests to us that, at least in principle, and if appropriately implemented from a computational standpoint, SW could be effectively employed in the future for creating, along the lines proposed by Russo (2005), a much needed reference lexicon of LIS based on corpora drawn from actual usage, and representing the important variation we know exists in LIS.

These promising indications certainly must still be carefully tested, and much more theoretical and empirical work is needed before we can say anything more conclusive. For example, the metalinguistic observations and discussions that have been stimulated, in our group of LIS signers, by the opportunity of "objectifying on paper" the forms of their language have suggested the need to explore more in depth the links, and distinctions, between written and "face-to-face" forms of language, and to achieve a clearer understanding of the similarities and differences between signed and spoken/written languages with respect to this dimension. The need of knowing much more on the history of writing in general, and of different writing systems has also arisen. We expect that comparisons beween written and transcribed texts of LIS and French Sign Language (LSF) we have planned to conduct in collaboration with our French colleagues will provide additional, valuable information (Garcia & Dalle, 2005).

From a broader socio-cultural perspective, we are obviously aware that, since writing is an inherently cultural process, the experimentation we have started within a very small group of LIS signers must be validated through a thourough confrontation with the larger Italian Deaf community. Whether a written form of LIS will or will not evolve will depend entirely from this community, and its cultural needs. What seems unquestionable to us is that whole issue of writing SL (as distinguished from transcribin and/or coding) needs to be considered much more carefully than it has been done thus far. This may open new, meaningful perspectives in our search for a clearer understanding of SL structures, and of more appropriate means for representing them.

4. References

Bébian, A. (1825). *Mimografie, ou Essai d'écriture mimique propre a régulariser le langage des sourd-muéts.* Paris: Louis Colas.

Bergman, B., Boyes Braem, P., Hanke, T., & Pizzuto (guest eds.) (2001). *Sign Language and Linguistics*, (*Special Issue on Sign Transcription and Database Storage of Sign Information*), 4: 1/2.

Biber, D., S. Johansson, G. Leech, S. Conrad, E. Finegan (1999). *Longman Grammar of spoken and written English*. London: Longman.

Brennan, M. (2001). Encoding and capturing productive morphology. In B. Bergman, P. Boyes Braem, T. Hanke & E. Pizzuto (guest eds.) (2001). *Sign Language and Linguistics*, (*Special Issue on Sign Transcription and Database Storage of Sign Information*), 4: 1/2, 47-62.

Cuxac, C. (2000). *La langue des signes française (LSF) ; les voies de l'iconicité.* Faits de Langues 15/16, Paris: Ophrys.

Fabbretti, D., & Pizzuto, E. (2000). Dalle mani alla carta: Aspetti teorici e metodologici della notazione della Lingua Italiana dei Segni. *Rassegna di Psicologia,* 2, vol. XVII, 101-122.

Fusellier-Souza, I. (2004). Linguistic variation and pragmatic aspects in sign languages considered from a semiogenetic point of view. Paper presented at the *International Colloquium "Verbal and Signed Languages: comparing structures, constructs and methodologies*". Rome, October 4-5, 2004

Garcia, B. & Dalle, J. (2005). Le projet français "LS-script" pour une formalization graphique des langues des signes, et l'experience SW en France. Seminar given at the *ISTC-CNR & ISSR, 2005-2006 Seminars' Cycle,* Rome, October 27, 2005.

Halliday, M. A. K. (1985). *Oral and written language*. Victoria, Australia: Deakin University.

Jouison, P. (1995). *Écrits sur la langue des signes française.* Paris: L'Harmattan.

Ong, W. J. (1982). *Orality and literacy.* London & New York: Mathuen.

Pennacchi, B., Pizzuto, E., & Rossini, P. (2001). La LIS: trascrizione e scrittura. Seminar given at the Institute of Psychology, CNR, Rome - *2000-2001 Seminars, CNR&The State Institute of the Deaf,* April 2001.

Pietrandrea, P. (2000). Sistemi di analisi e rappresentazione delle lingue dei segni. Seminar given at the Institute of Psychology, CNR, Rome - *1999-2000 Seminars, CNR & State Institute of the Deaf*, February 2000.

Pizzuto, E. & Cuxac, C. (2004). *Language, its formal properties and cognition: What can be learned from signed languages*. Joint Italian/French CNR/CNRS Project (ISTC-CNR & Université Paris 8 / CNRS Unit UMR-7023 - "Structures Formelles du Langage").

Pizzuto, E., Pennacchi, B., Rossini, P., Sutton, V. , Pietrandrea, P., Russo, T. (2000a). An Issues- Raising Approach applied to different written representations of a single LIS utterance. Paper presented at the Workshop "Cross-linguistic investigations of signed languages: can similarities and differences be detected without appropriate tools for representing and analyzing signed texts?". *7th International Congress on Theoretical Issues in Sign Language Research*. Amsterdam, July 23-27, 2000.

Pizzuto, E. , S. Wilcox, T. Hanke, T. Janzen, J.A. Kegl, J. Shepard-Kegl (2000b). "Cross-linguistic investigations of signed languages: can similarities and differences be detected without appropriate tools for representing and analyzing signed texts?", Workshop presented at the *7th International Congress on Theoretical Issues in Sign Language Research*. Amsterdam, The Netherlands, July 23-27, 2000.

Pizzuto, E., & Pietrandrea, P. (2001). The notation of signed texts: open questions and indications for further research, *Sign Language and Linguistics*, (*Special Issue on Sign Transcription and Database Storage of Sign Information*), 4: 1/2, pp. 29-43.

Pizzuto, E, Pennacchi, B. e Rossini, P. (to appear), Representing signs in written form: issues to be addressed. *Proceedings of the International Congress on Aprendizaje de las Lenguas Europeas en personas Sordas y Sordiciegas en el siglo XXI - APREL XXI,* Barcelona, 30 November - 1 December 2001.

Prillwitz, S., Leven, R., Zienert, H., Hanke, T., Henning, J., et al. (1989). *Hamburg Notation System for Sign Languages. An introductory guide*. Hamburg: Signum Verlag.

Pulleyblank, D. (1987). Yoruba. In B. Comrie (ed.), *The World's Major Languages.* London: Crom Helm, 971-990.

Russo, T. (2000). *Immagini e metafore nelle lingue parlate e segnate. Modelli semiotici e applicazioni alla LIS (Lingua Italiana dei Segni)*. Ph.D. Dissertation, Universities of Palermo, Calabria and Rome "La Sapienza".

Russo, T. (2004). *La mappa poggiata sull'isola. Iconicità e metafora nelle lingue dei segni e nele lingue vocali.* Rende: Università della Calabria.

Russo, T. (2005). Un lessico di frequenza della LIS. In T. De Mauro & I Chiari (eds.), *Parole e numeri.* Rome: Aracne, 277-290.

Sallandre, M-A. (2003). *Les unités du discours en Langue des Signes Française. Tentative de catégorisation dans le cadre d'une grammaire de l'iconicit*é. Ph.D. Dissertation, Paris, Université Paris 8.

Stokoe, W. (1960). Sign language structure: an outline of the visual communication system of the American deaf. *Studies in Linguistics, Occasional Paper* (8) (2nd printing 1993: Burtonsville, MD: Linstok Press).

Streiter, M. & Rocha Costa (eds.) (2004). *Proceedings of the Workshop on the Representation and Processing of Sign Languages. LREC 2004* [online publication: www.signwriting.org/forums/software/lisbon].

Sutton, V. (1999). *Lessons in SignWriting. Textbook & workbook* (2nd edition, 1st edition 1995). La Jolla, CA: Deaf Action Commitee for Sign Writing.

5. Acknowledgements

We gratefully acknowledge partial financial support from the Istituto Statale Sordi di Roma (ISSR) and from CNR and CNRS (French/Italian Project "Language, its formal properties and cognition: what can be learned from signed languages" - ISTC-CNR, Rome & UMR 7023 CNRS, Paris).

A Sequential Approach to Lexical Sign Description

Michael Filhol, Annelies Braffort
LIMSI/CNRS
BP 133, F-91403 Orsay cedex
E-mail: michael.filhol@limsi.fr, annelies.braffort@limsi.fr

Abstract

Sign description systems able precisely to detail how a lexical unit of a sign language is performed are not that numerous. Plus, in the prospect of implementing such a description model for automatic sign generation by virtual characters, visual notation systems such as SignWriting, however accurate they are, cannot be used. The Hamburg Notation System (HamNoSys) (Hanke, 1989) together with its more computer-friendly super-set SiGML (Signing Gesture Markup Language) is about as advanced a model we could find, and yet some problems still have to be tackled in order to obtain an appropriate sign description system. Indeed, based on Stokoe-type parameters, it assumes every sign can/must be described with the same fixed set of parameters, each of which would be given a discrete value. However, we argue that not all signs require all parameters, and that not all the parameters that are needed can be given at the same time in the same way. This work underlines three problems we see with Stokoe-like descriptions, and suggests a new approach to handling sign language lexicon description.

1. Over-Specification

The trouble when filling all parameters with values is that they all inherit the same status. Yet often, some are crucial to the sign in that changing them would destroy the whole sign, whereas others are only given so as to enable, say, a signing avatar to perform the target sign but could well be specified differently. For instance, the palms of both hands in the sign [WHAT]$_{LSF}$ need be horizontal and facing up, but the fingers may point to anywhere away the signer's body (fig. 1). Actually, the direction they point to may even vary through time, as signers usually prefer to rotate around the wrist or elbow rather than around the shoulder. With a HamNoSys notation, both "fingext" orientations *out* and *out-left* (for a strong hand on the right-hand side) would define the [WHAT]$_{LSF}$ sign properly, but one has to be chosen.

Figure 1: [WHAT]$_{LSF}$ (Moody, 1986)

The recent addition of the "..." subscript operator in HamNoSys v4 allows to "soften" a value and change it to a somewhat fuzzier specification. That is, used with our example, turn the fingext *out* value into something like "*out* or *out-right* or *out-left*". However, nothing precisely defines this operator, and applying it to the fingext *out* value will also make valid values like *out-up* and *out-down*, which we do not want.

The source of the problem above is that the fingext direction was "hard-wired" to a particular value, and then softened. Instead of over-specifying and merely stating what can be approximated, we suggest that the sign contents should be constrained enough to define the target sign, but that whatever is not necessary be banned from its description. On our example, setting the palm plane normal to an upright vector is enough about the hand's orientation in [WHAT]$_{LSF}$.

2. Parameter Dependencies

Secondly, parameter models consider the parameters separately and each of them is assigned a distinct value, with no regard for other parameters. In computer science terms, none of these assignments is in the scope of another, so each of them could be carried out in a parallel way, i.e. all at once and independently. Though, this does not account for inter-parameter dependencies, such as that in [DESK]$_{LSF}$ (fig. 2). The strong hand movement depends on the fingext direction (in HNS terms) of the weak hand, whichever is chosen in the first place.

Figure 2: [DESK]$_{LSF}$ (Moody, 1986)

The issue of parameter dependencies was already addressed and partly resolved with the new HamNoSys "~" subscript operator. It is applicable to palm orientation or fingext direction to make it relative to the path of the corresponding hand. It allows descriptions such as that of $[\text{BALL}]_{\text{DGS}}$, making palm orientation relative to its path on each hand. In $[\text{DESK}]_{\text{LSF}}$ however, the dependency is not one of a hand orientation on its path, but that of one hand's path on the other's orientation.

Moreover, two different parameters could well depend on a common non-parameter object, such as in $[\text{BUILDING}]_{\text{LSF}}$ (fig. 3). The strong hand moves along and close to a line, say *L*. Its palm is constantly facing *L* and the weak hand's location and orientation is defined as being symmetric to those of the strong hand's, with respect to *L*. Both location and palm orientation of both hands depend on the same line *L*. Although *L* is obviously crucial to the sign as a great part of the description depends on it, no parameter (in Stokoe's sense) is ever *equal* to *L*, which is why we call *L* a non-parameter common dependency.

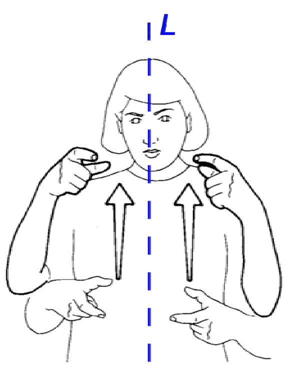


Figure 3: $[\text{BUILDING}]_{\text{LSF}}$ (Moody, 1986)

To account for the two cases stated above, we claim that any part of a sign description should be allowed to make use of other parts, even of the same description. This way, internal dependencies become part of the description.

3. Iconic Structures

Above all, using C. Cuxac's theory (2000) of iconicity as a framework for ours, it has become obvious that the many possible context influences cannot be ignored while modelling lexical description. A great part of sign languages' beauty and power in concision comes from the potential for signs to be altered according to the context in which they are used, thereby switching discourse from a conventional sign flow to *highly iconic structures* (HISs). For instance, the sole sign $[\text{BOX}]_{\text{LSF}}$ can be used to sign the phrase "large box" in LSF, only the distances between the hands will be greater than the ones involved in the plain conventional $[\text{BOX}]_{\text{LSF}}$ sign (plus the signer will probably also puff his cheeks and raise his elbows).

There are many forms of iconicity in SLs: size&shape transfers, personal/situational transfers, use of time lines... Formalising such features for automatic sign generation is not trivial. Some work has been initiated with the ViSiCAST project to include use of proforms and signing space in particular (Hanke *et al*, 2002), but we found nothing close to the richness emphasised in (Cuxac, 2000). An HIS can not only alter the location or the hand shape involved in a sign, but also a path, a direction, eye gaze, etc. Virtually, anything can be acted upon, and these actions being commonplace in SL, we claim a description model should allow signs to behave accordingly. Back to the example above, describing $[\text{BOX}]_{\text{LSF}}$ without making the distance between the hands responsive to the contextual size weakens the sign's re-usability.

4. A Geometrical Approach to Descriptions

We are now ready to outline a proposal for a new sign description model whose aim is to make for the three main problems we see with present parametric models, stated above and summarized below :

- unnecessary parts should not appear in a description;
- the different parts should be able to refer to one another;
- descriptions should be made flexible enough to be responsive to context influences.

Specifying What is Needed and Allowing Internal Dependencies

We handle the first two points using a statement-based language, each of which is either a build statement (B-statement) or a constraint statement (C-statement). B-statements are used to build objects like points, vectors or planes that can be referred to in subsequent statements. C-statements serve the main point: a C-statement either assigns a value to an existing object or adds a constraint to it. Constraints are either applied to an object itself or to one of its "slots" if it has any. A slot is a constituent of an object that accepts geometrical constraints but may remain unmentioned. For example, eyebrows can be set to frown in a sign *S* by slotting a value in the appropriate slot, denoted `S.eyebrows`. Yet in other signs the eyebrows can stay unspecified, and indeed they often do.

The syntax used for C-statements is close to that used in mathematical definitions of geometrical figures (Filhol, 2006). For example, the following C-statement sets a correct orientation (the `ori` slot of the hand) for the strong hand (the `shand` slot of the sign) of a sign `S` by constraining its palm (the `palm` slot of the orientation) to be orthogonal (the "`_|_`" operator) to a direction pointing up (`Up` is a constant) :

```
S.shand.ori.palm _|_ Up
```

The syntax used for B-statements resembles that of variable declarations in most programming languages, i.e. a type keyword and an identifier. Here is a B-statement that creates a plane named `P`:

```
PLANE P
```

Most probably, C-statements enrolling `P` will follow, in order to constrain it and use it afterwards as an internal dependency. For example, to make it horizontal:

```
P _|_ Up
```

Note: Parsing such an input will require some conflict-checking, as two contradictory C-statements applied to an object should be rejected. Though we shall not deal with this issue here.

With this description language, sign descriptions can be specified as much – or indeed as little – as wanted, which tackles the first drawback underlined in part 1. Moreover, each part of a sign description can make use of any other part, provided the latter has been defined beforehand. Thus, contrary to parametric models, value assignments are no more paralleled but made sequential, and values are not only chosen from a fixed set but may depend on intermediate objects (there again provided they were built earlier on) if any are needed. An acyclic dependency graph can then be associated with the description, which represents the description's internal dependencies.

Iconicity in Descriptions

Although no implementation has been done on this issue so far, it has always been regarded as a necessary prospect in the design of our description model. Here is how we will extend the given language to handling iconicity in sign descriptions.

Enabling iconicity in signs can be done by extending the language with a new type of reference. Every time a value or an object is expected in a statement, a call to a context element can be placed instead. For instance, instead of specifying an arbitrary distance between the hands' positions in the description for $[\text{BOX}]_{\text{LSF}}$, we may refer to an external reference called *size*. This way, whenever the description is used to perform the sign in discourse (i.e. in context), it can be tagged with a size attribute, so that the distances are altered accordingly, with no extra rule about how to sign "big box" or "small box".

This brings us to extend the dependency graph to external nodes, in the sense that some of the values within the description will depend on values that are "outside" the lexeme itself. In fact, they are to be found in (or given by) the context/syntactic level.

More generally speaking, this comes down to including semantic information in the lexical units being described. Indeed, it is a reasonable hypothesis that the list of external dependencies relates the cognitive type of the sign's concept. E.g. $[\text{BUILDING}]_{\text{LSF}}$ will at least have the following external dependencies : height and width and situation in signing space. The results we have started to collect from our study of the French conventional lexicon go to show that a lot of signs denoting concrete objects have the same physical dependencies, namely size and location.

5. Full Example

Here is a full example of a description for $[\text{BUILDING}]_{\text{LSF}}$, drawn in fig. 3 further up. Figure 4 illustrates the various objects built within. External dependencies labels are between curly brackets; the outfix `|x|` operation stands for the length of the argument vector `x`; infix `/\` is the vector product operator.

```
1.  SIGN S

2.  LINE L
3.  L // Up
4.  L THRU {Loc}

5.  POINT M
6.  VECTOR V
7.  V = Vect({Loc}, M)
8.  V _|_ L
9.  |V| = {Size}

10. S.shand.config = "BSL C"
11. S.shand.ori.palm = -V
12. S.shand.ori.fingext = Up /\ V
13. S.shand.traj.start = M
14. S.shand.traj.mvt = {Height}*Up

15. S.whand SYM S.shand WRT L

16. REGISTER S "building"
```

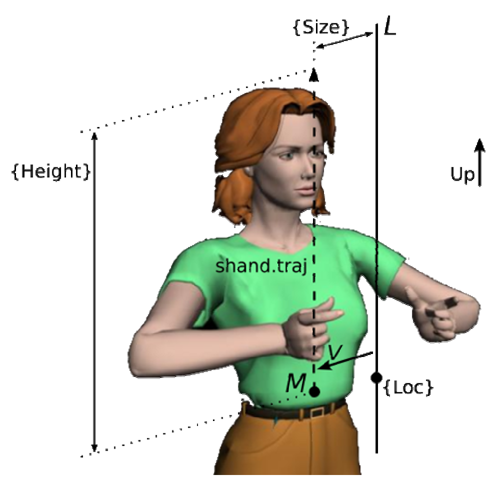


Figure 4: Objects involved in description of $[\text{BUILDING}]_{\text{LSF}}$ below

Line 15 indicates that the weak hand must be symmetric to the strong hand with respect to line *L*. We give it here as an example of the type of C-statement the model might end up with. It actually means that:

- configurations are identical;
- locations verify the given symmetry;
- palm fingext vectors are identical;
- palm normal vectors verify the symmetry.

Hence, line 15 really is a short for:

```
15a. S.whand.config =
           S.shand.config
15b. S.whand.loc SYM
           S.shand.loc WRT L
15c. S.whand.ori.fingext =
           S.shand.ori.fingext
15d. S.whand.ori.palm = V
```

However, we are not yet able to tell whether hand symmetries all behave this way, whatever the sign being described. The only genuine symmetry related in this statement is the one that applies to the hand locations (see line 15b). It may indeed turn out, say, that both hands of a two-hand sign where locations are symmetric along a line have the same normal vector. Please note that the description language is still under development.

6. Conclusion

What we have outlined here is a new way of addressing the description of sign language lexicon units. Instead of merely giving independent values to a given set of parameters, it is based on sequences of constraint statements, which unlike previous models make use of internal dependencies between the elements of the descriptions. Consequently, all the units described do not necessarily mention the same information, but rather each description only states what is needed.

To assess this **geometrical and sequential** approach, we are planning on describing signs on a larger scale. We believe that the flexibility of the suggested language itself will make it easy to cope with many types of constraints, if more are needed. A practical concern in the design of this model is also to limit the number of possible descriptions for a given aspect of a sign, as the fewer there is, the more sign descriptions will look alike, and the more useful the model becomes as to categorize the signs with respect to their descriptions' (or their dependency graphs') layout.

7. References

Braffort, A. (1996). Reconnaissance et compréhension de gestes – Application à la langue des signes. Thèse de doctorat. Université d'Orsay.

Cuxac, C. (2000). La Langue des Signes Française – Les voies de l'iconicité. Faits de Langue n°15-16. Ophrys.

Filhol, M. (2006). Une approche géométrique pour une modélisation des lexiques en langues signées. In proceedings of TALN-Recital 2006, to appear.

Hanke, T. (1989). HamNoSys – An introductory guide. Signum Press, Hamburg.

Hanke T. *et al* (2002). ViSiCAST deliverable D5-1: interface definitions. ViSiCAST project report. `http://www.visicast.co.uk`

Hanke, T. (2004). Towards Sign Language resources – Extending an HPSG lexicon for German Sign Language from empirical data. TISLR 8, Barcelona.

Moody, B. (1986). La Langue des Signes – Dictionnaire bilingue élémentaire. IVT, Paris.

Italian Sign Language (LIS): can we write it and transcribe it with SignWriting?

Alessio Di Renzo[1], **Luca Lamano**[1], **Tommaso Lucioli**[1], **Barbara Pennacchi**[2], **Luca Ponzo** [1]

[1] Istituto Statale Sordi di Roma (ISSR)
[2] Istituto di Scienze e Tecnologie della Cognizione – CNR (ISTC-CNR)
Via Nomentana 56, 00161 Roma (RM), Italia
Alessio Di Renzo <alexiodirenzo@tiscali.it>, Luca Lamano <vegeta_luca@libero.it>, Tommaso Lucioli <tlucioli@inwind.it>, Barbara Pennacchi <barbara.pennacchi@istc.cnr.it>, Luca Ponzo <kikko9@hotmail.com>

Abstract: This paper is the result of our discussions and reflections on using Signwriting for writing and transcribing LIS texts

1. Introduction

The main aim of this paper is to present and discuss some of the most relevants issues/points arising from our direct experience (as Deaf LIS signers) with the problems of representing signs. Our research has been developed within a joint ISSR/ISTC-CNR project on “Writing LIS and SignWriting”. The project started in February 2005, and aimed to provide an opportunity for a small group of Deaf signers (already involved in sign language research and Deaf education) to reflect together on different notation systems that have been devised for signed languages (hereafter: SL), and to explore more in depth the possibility of using Sutton’s (1995) SignWriting system (SW) for representing LIS texts in a written form that could satisfy our needs more appropriately than other notations we had previously used or explored in transcribing and analysing SL data.

Our research is grounded in previous work conducted within our group on several theoretical, methodological and practical problems arising from the fact that, to date, no SL has spontaneously evolved a written form, and there is still no general consensus on what are the most adequate tools for writing and/or transcribing SL (Fabbretti & Pizzuto, 2000; Pietrandrea, 2000; Pizzuto & al, 2000, and to appear; Bergman et al, 2001; Pennacchi et al, 2001; Pizzuto & Pietrandrea, 2001). Our work is also related to relevant crosslinguistic work on notation and representation issues carried out within an ongoing broader French/Italian crosslinguistic project on LIS and French Sign Language (LSF) (Pizzuto & Cuxac, 2004, Garcia & Dalle, 2005).

The main objectives of the work we are doing with SW are:

1) explore the adaptability/feasibility of SW as a system for transcribing/coding SL texts accurately, without using the “pseudo-standard” system of pictures and “glosses”;
2) explore its usability as a LIS writing system, although it still remains to be seen whether and/or when the Italian Deaf community would adopt it as such.

The reason for choosing to explore SW instead of other notation systems is due to the fact that almost all the other notation systems (such as HamNoSys [Prillwitz & al, 1989], to cite one of them) are either Stokoe-based or focus mainly on describing in detail single signs. When using these systems with streams of signs tightly linked to each other as in a SL discourse or dialogue, notation becomes rapidly a cumbersome affair. In addition, none of these notation systems can be easily used by Deaf people as a writing system for expressing themselves in their own language.

2. “Oral” languages and writing systems

If one tries to consider the wide diversity of languages used in the world, taking in due account the fact that the largest majority of them (around 90%) do not have a writing system, and that there is also a wide variety of writing systems (Ong, 1982; Breton, 2003), the task of reflecting over the relation between an “oral” language and its writing system may seem daunting at first.

But there are some common elements that mark the difference between languages with and without written systems. The social relevance of being able to spread and preserve information through space and time is one. The fact that writing has enabled people to keep track of their language's evolution and variation through time and space is another. Another thing to note is that there isn't yet a writing system that could actually display *all* the elements of human speech. And not all writing systems are purely phonologically-based. But all serve the same purpose: to enable a reader to “rebuild” in its mind, or read aloud, what the writer wrote.

However, so far, all writing systems in use today are for languages that use voice and sound. SL are still without a writing system. This makes a large portion of Deaf people live in a diglossic environment, where they're forced to use one language (their SL) in their face-to-face interactions and another language in all other types of human interaction. And the well known fact that most Deaf people have a lower proficiency in the written language of their country, in comparison to their SL skills, renders the situation very complex from a socio-cultural point of view.

This is one of the reasons that made us try out SW, in the hope it could be a good candidate for becoming a writing system for SL, as it is structured in such a way that it can be written by hand or on a computer (by using specifically designed software), with a consistent set of visual rules that are easy to memorize.

3. Writing and transcribing LIS texts

Soon after beginning to learn and discuss the various SW glyphs[1] and their adaptability to LIS signs, we felt the need to explore the use of SW for both creating written LIS texts, conceived and expressed from the start in a written form (something we had never tried before), and for transcribing LIS narratives originally produced in the “face-to-face” modality that is prototypical of all SL. We

[1] We use this term to refer to any individual graphic element belonging to SW’s set of symbols. We feel that this term is more appropriate than other definitions more semantically loaded, such as ‘graphemes’, ‘characters’ or ‘symbols’.

use here the term "face-to-face" (for which we have a sign in LIS) to characterize the visual-manual form of signing, analogous to the "oral" form of spoken languages.

Thus far we have produced four written texts, two transcriptions of portions of signed texts, and one translation from written Italian to written LIS of a classic tale by Aesopus. Some of these texts (all handwritten) are very short (from 8 to 14 sign/graphic units), others longer (from 31 to 57 units). The first two texts were produced, though with some "ortographic" errors, after only 6 hours of learning the basics of SW. This in itself is interesting: it indicates that the learning curve may be less steeper than one could imagine 'a priori', at least for Deaf people. We also found from the start that when texts produced by one of us were read by different signers (not just by their author), the readers were able to accurately "rebuild" (e.g.: to sign "aloud") the signs encoded in SW glyphs, and interpret the overall meaning of the texts, in a way that we have never experienced with any other notation for SL. On this basis, the ensuing discussions focused on different problems, including those related to the 'orthographic' choices to be made.

For example, one such problem concerned the left-to-right vs. top-to-bottom ordering of the sign units in the text. After trying both orderings, we found that we preferred a 'top to bottom' order. This ordering helped us to represent more clearly spatial modifications of the signs (e.g. lateral shifts in space) that convey important grammatical information in LIS.

The most interesting result we obtained from the start, however, was that, compared to other systems, the SW glyphs could be used to represent LIS signs in a way that was easier, richer, and much more efficient for signers. Most importantly for us, the SW glyphs appeared to allow us to represent relevant structural features of the visual-spatial lexicon and grammar of LIS.

3.1 Writing "standard", "non-standard" signs and relevant nonmanual components

With the SW glyphs we were able to write down not just "standard signs" that are listed in the available LIS dictionaries, but also complex signed constructions (of equally complex meaning) that are very frequent in signed discourse, yet are not listed or are just mentioned in LIS dictionaries and grammar as part of the "non-standard" or "productive" lexicon. These constructions, which we currently consider as Highly Iconic Structures (HIS) within the frame proposed by Cuxac (2000), include different types of manual and nonmanual elements that are described in the literature with various terms such as "classifiers", "role playing" or "impersonation" devices. (see Emmorey, 2003, Pizzuto & al, 2005; Russo, 2004). The two-sign sequence in Figure 1, taken from an early LIS text ("At Home") written by TL in a left-to-right order, provides one example.

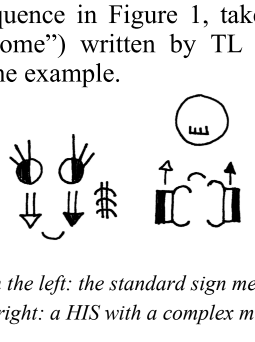


Figure 1 - On the left: the standard sign meaning 'snow' on the right: a HIS with a complex meaning

In Fig. 1, the standard sign for 'snow' is followed by a HIS unit: a non-standard manual sign, with a very specific facial expression, translatable as *"a really thick coat"*. Within the context of the LIS text, the two-sign sequence means "there was a really thick coat of snow".

A feature that struck us immediately as we reflected on our written texts was our own 'spontaneous' use/non use of the SW glyphs for the nonmanual components of signs. This made us more aware of regularities in the LIS lexicon and grammar that we had not been able to detect previously (see also section 4). For example, the standard sign on the left in Figure 2 (from TL's text mentioned above), means 'stuck', and was written with a specific mouth-glyph ('half-protruted mouth'). After discussing, we discovered that this allowed us to differentiate this sign from a related one (on the right in Fig. 2) with a different mouth-glyph ('showing-teeth'), and an equally different meaning.

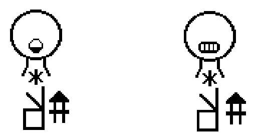


Figure 2 - Two different standard signs for "stuck"

The difference in meaning between these two standard signs appear to be broadly related to the expression of modality in LIS (Wilcox & Wilcox, 1995): while the first sign means "being stuck, and accepting this state of affairs, without any possibility to change it ", the second one means "being stuck but with the necessity, urgency and possibility to do some action directed to modify this state of affairs".

We uncovered, in other texts, other cases in which a different mouth gesture vehiculates differences in meaning between standard signs that have the same manual form. Another example, enlightening also for its variability across signers, is given below, in Figure 3.

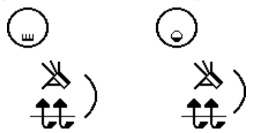


Figure 3 - Other two signs, both meaning "very special"

Both signs mean "very special". Sign (a), on the left, is taken from a text written by LL; the other sign (b) was identified following discussions over the mouth gestures that appear to be an integral part of these signs: an 'upper teeth over lower lip' for (a), and 'half-protruted tongue' for (b). For some signers the two sign variants can be used interchangeably, whereas for other signers they cannot: variant (a) must be used when the "very special" quality attributed to something is based on somebody's internal judgement, while (b) is required when the same quality is 'more objective', stemming from the object itself.

The point of these observations is that the analysis of different form-meaning correspondencies, as it could be accomplished via written representations of LIS texts, provided valuable indications on the relevant manual and nonmanual components of the signs interconnected in a text, and their stability and/or variability across signers (this variability is to be expected, given the lack of a written tradition in LIS).

3.2 Representing morphological modifications and discourse relations

Figure 4 shows a more complex sequence, extracted from a text written by LL (top-to-bottom order), in which the author describes how, on a Christmas-vacation day, his father woke him up to ask him to go together to the father's home village. The fragment reported below describes a direct-discourse interaction between the author and his father, and could be translated as follows: *"(...) I woke up reluctantly and, from under the blankets, asked him "what do you want?" He said "let's go, the two of us, to my home village (...)".*

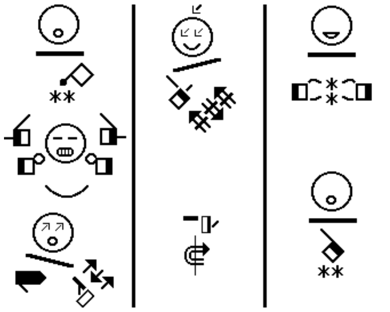


Figure 4 - Excerpt from a text written by LL

We found especially valuable the way in which the SW glyphs allowed the author to represent, and the readers to successfully rebuild, structural features that are unique to the signed modality. For example, the glyphs for the manual and nonmanual components of the second sign in the first column accurately represent the alterations of the movements of the hands, and of the facial expression which mark a morphological, aspectual modification of the base sign for 'wake-up' to vehiculate the meaning 'wake-up-reluctantly'.

Even more interesting for us was to find out how effectively the SW could represent another kind of HIS typical of LIS (and more generally SL) face-to-face discourse. These are usually described by signers as "impersonation" devices because, via changes in gaze directions and postural modifications of the shoulders or upper trunk orientation, the signer appears to "impersonate" the referents he is reporting about, or whose utterances he is quoting, as in the fragment described in Figure 4.

In the third and fourth major graphic units in Fig. 4 these impersonation markers are represented by the eye-glyphs encoding 'eye gaze up diagonally' and 'eye gaze down diagonally' (the two arrows within the circles in the third and fourth unit, respectively), together with the shoulder-glyphs encoding congruent 'shoulder orientation modifications' (the horizontal 'bars' oriented upwards and downwards in the same units three and four).

These glyphs are superimposed on those for the manual signs: the resulting "gestalt" of spatial disposition encodes very clearly the structural links between the manual and nonmanual components because it mirrors how, in actual signed discourse, nonmanual impersonation devices are temporally superimposed on manual signs, and distinguish the referents they identify. In this case, the impersonation marker in the third written unit identifies the writer/author, while the one in the fourth written unit identifies the author's father, both referents being represented in a "first person role". Due to space constraints, our considerations will be limited to the glyphs that compose the third complex unit of this written text.

The impersonation mark encoding the writer in a first person role is superimposed on the glyphs for two manual signs, meaning, respectively: 'staying under the blankets' (on the left) and 'what do you want?' (on the right). The spatial disposition of these two written signs, one next to the other, also shows that they 'co-occur in space'. This spatial arrangement of the written units mirrors the spatial-temporal arrangement the corresponding manual signs may have in actual signing, where they could be either simultaneusly co-articulated or one sign could be maintained in space and time while producing the other, i.e. articulating with the left hand the sign written on the left, and with the right hand the sign on the right. In fact this is how the written text was read and signed "aloud" by readers other than the author.

Another thing to note about the "gestalt" of the written signs under discussion, is the mouth-glyph (the small circle whithin the wider circle of the 'face-glyph'). We found that this mouth-glyph was necessary to distinguish the hand-glyph on the left ("what do you want?") from an almost "homographic" glyph for a partially (semantically) related LIS sign meaning "why?". In their signed, face-to-face forms, these two LIS signs have the same handshape and movement, but different mouth gestures (see Fig. 5 below), and this distinction was quite naturally signalled in the written rendition of the two signs.

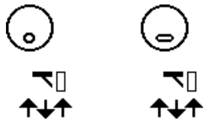


Figure 5 - On the left: the standard sign for "what do you want?" On the right: the standard sign for "why?"

3.3 Writing vs. transcribing

A relevant outcome of our work has been a much more thorough, empirically grounded understanding of the important differences between 'writing' and 'transcribing'.

We realized that, when writing, choosing the glyphs to represent what we meant was relatively simple: we put down on paper only those "articulatory gestures" that, relying on implicit intuitions, we believe we make when producing signs meaningfully structured in discourse (e.g. see example in Fig. 4). Then we "tried out" the efficacy with which our written texts conveyed what we meant by having others read them.

Obviously, since we are not used to create written LIS texts, in some cases our writing was somewhat too close to the 'face-to-face' LIS form, and some ambiguities arose. For example, in one case, the written text did not provide sufficient information to identify which of two characters of a narrative performed a given action. But the ambiguities we spotted appeared to be on the whole comparable to those that may be found in texts written by vocal language speakers who are not very familiar with the written modality of language expression, hence produce texts that are closer to an "oral" form of language, where information that is necessary in writing can often be omitted without compromising speakers' comprehension.

When producing transcriptions, clearly we could not rely on our own intuitions on how signs are made. We had to try to transpose on paper, as accurately as possible, all the articulatory gestures that we felt were meaningful for subsequently "rebuilding" and analysing the original signed performance. But this objective poses many more problems that one can think of beforehand. We will briefly illustrate here only the most general and rather obvious one: the need of deciding what exactly is relevant, for producing an accurate transcription, and what can be left out.

The example in Figure 6 was excerpted from the first version of a transcription, made by LL, of a text in which a signer reported on "four monkeys escaping from their cage". The short sequence in Fig. 6 represents: (a) in the left column, two signs meaning 'cage', marked at two locations in space to mean that 'there were two cages'; (b) in the right column, three signs meaning 'closed', also marked at three different locations in space to mean that 'each of three cages [referred to] was closed'.

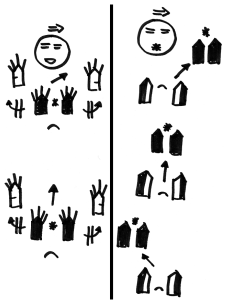


Figure 6 - Excerpt from a transcription made by LL

The transcription revealed that the original signed text contained an 'error': the sign for 'cage' should have been produced three times instead of only two times, because the 'cages' referred to were three, not two. But we wish to note here also another aspect relevant for understanding the problems we faced. Upon reading, the glpyhs allowed us to "recover" on our own some important nonmanual aspects that we knew must have been in the original signed text, but didn't appear in the transcript. Thus a discussion arose as to whether the transcription was accurate and consistent, especially with respect to nonmanual signals.

We checked the original signed version, and we found that *each* dislocation in space of the manual signs occurred with congruent nonmanual markers (shoulder, eye-gaze and head displacements) which, however, the transcription represented only partially (e.g. by a head-displacement glyph, annotated only over the first sign for 'cage' and the first for 'closed'). The displacements of the manual signs were also transcribed somewhat differently: only via arrow-glyphs for 'cage' vs. arrow-glyphs plus a different collocation on the page for 'closed'. These observations led us to revise the transcription, adding a more complete description of nonmanuals and spatial dislocations.

We noticed also that, when comparing transcriptions with written texts, the SW transcripts tend to contain more facial glyphs that aren't strictly related to the content of the narrative, such as prosodic expressions, like hesitations or "pauses of reflection", while in the written texts we produced this kind of prosodic glyphs are absent. This detail made us even more aware of the conceptual and empirical differences between transcribing and writing.

This type of problems are largely comparable to those found in transcribing spoken language data. As Ochs (1979) has clearly shown with respect to spoken texts, transcription is a theory, and deciding what needs to be selected or not to be written down, and how to annotate it for producing an appropriate transcription is a very complex task, highly dependent from the specific objectives pursued in equally specific investigations. Both the objectives pursued and the criteria adopted for transcribing must be made explicit and motivated on theoretical grounds. This task is difficult in research on spoken languages, and clearly even more difficult in research on SL, where the absence of a written tradition renders everything more problematic.

4. Writing decontextualized signs

As we proceeded in our work with LIS texts, we realized that we needed to do a complete adaptation of Sutton's (1999) SW manual for use within the Italian Deaf community. When we started, we relied upon a partial adaptation of the manual, including an Italian translation of the English text (realized by Cecco [2001]), but illustrative examples were still based on American Sign Language (ASL). A clear understanding of how to use the SW glyphs thus required knowledge of ASL signs, which some of us had, but others did not. In order to use the SW manual more productively among ourselves, and also for making it accessible and usable outside of our small group, within the broad community of LIS signers, we needed to illustrate the SW glyphs with appropriate examples based on LIS, not on ASL.

At first, this task seemed simple enough: we thought we would just look for LIS signs that would be adequate substitutes for the original ASL signs. But, when we started working on this, we found out that there were many other issues to deal with.

For example the fact that a sign can be written in more than one way, depending on what level of detail one desires to convey, and on the fact that the reader must still be able to understand it without being overwhelmed by information overload.

Or the fact that ASL and LIS present differences in the frequency of usage of different hand configurations. SW's set of hand-glyphs includes all handshapes that a human being could make, but each SL has different handshape usage frequencies (Volterra, 1987/2004). However, at least for LIS, these frequencies of usage have been extrapolated from LIS dictionaries (Pietrandrea, 1997; Radutzky, 1997). Unfortunately, in our opinion, these dictionaries are based on the flawed assumption that the citation form of a sign would also be the most used within "face-to-face" LIS communication. We think that, in order to produce more reliable LIS dictionaries (i.e.: more descriptive of real LIS usage), it is necessary to analyze also "real" signs, such as one might find within a SL text, either written or "face-to-face" (and then transcribed).

While hunting for LIS examples to use in the adaptation of Sutton's SW manual, we have collected and written down about 600 single signs which we have, in some sense, extracted from our 'mental lexicon'. It has been quite natural for us to reflect on similarities and differences between the ways in which we have represented these decontextualized signs, compared to the signs occurring within our written and transcribed texts. We mention here only two of the major similarities and differences we have noted.

First, almost all of the decontextualized signs we have written for illustrating the SW glyphs appear to belong to the class of “standard” signs, while very few belong to the class of HIS. This seems to us particularly interesting if we think that the use of HIS *is* very common in actual signed discourse. It indicates us two things: (a) that decontextualized signs alone cannot be used as the only or primary source of informations on the LIS lexicon; (b) that HIS signs are, by their nature, highly interconnected with their context of usage and cannot be decontextualized without some “semantic damage”. In our opinion, this means that, if we want to have in some future really accurate LIS dictionaries, we have to revise their present structure and procedures for collecting lexical items.

Second, there were marked differences in the way we used glyphs for meaningful nonmanual signals, especially facial expressions, when writing decontextualized signs vs. text-framed signs. In general, most decontextualized signs appeared to not require nonmanual glyphs, while for most signs framed within a text we felt that nonmanuals were necessary components to be written down.

These impressions were supported by a preliminary analysis we made by comparing all the LIS sign units within our written texts and transcriptions (232 units), with an equivalent number of decontextualized LIS signs taken from our adaptation of the SW manual. We found that 70% of text-framed units were written with glyphs for meaningful facial/gaze/mouth/postural gestures (in addition to the glyphs for the manual parts), while the remaining 30% showed only the signs’ manual components. This distribution was reversed in decontextualized signs: the vast majority (75%) were represented with glyphs for only the manual components, while a markedly smaller proportion (25%) included also glyphs for nonmanuals.

5. Some indications for further research

Our project is still ongoing. We have almost completeted the LIS/Italian adaptation of the SW manual, and we are producing more written texts and transcriptions. However, the corpus of texts and individual signs we have assembled thus far is certainly not enough to evaluate to what extent SW will prove to be a valuable tool for both writing and transcribing LIS.

We need to collect and analyze more texts written directly in LIS, and more transcriptions of different genres (e.g. monologues, dialogues, free and elicitated narratives, poetry, texts produced during lectures or of ‘explicative’ rather than narrative type). We have planned relevant crosslinguistic comparisons between LIS and LSF data.

We want also to broaden our reflections on writing systems in general, as this can certainly help us in our search for the best way to write down our language.

The analyses we want to conduct require the creation of databases, and the improvement/development of computational tools. We plan to use SignPuddle (http://www.signbank.org/signpuddle), with appropriate implementations as needed for LIS data. Currently, there are some attempts to include SW glyphs within Unicode, the Universal Character Encoding containing all different graphemes of almost all world's written languages. The inclusion of SW glyphs in Unicode may well ease considerably the creation and the use of present/future databases and writing and/or research software (see http://www.signwriting.org/archive/docs1/sw0037-SW-In-Unicode.pdf and Aznar, G. & Dalle, P. in this volume).

While much remains to be done before saying anything more conclusive, the results obtained so far provide some relevant indications with respect to: (1) the representation of signed language data; (2) corpus collection and construction for signed languages (at the lexical and textual levels).

With respect to corpus collection and construction, our work suggest that it is very important to focus from the start on the problems posed by text corpora, rather than focusing only on corpora built from annotating/eliciting individual lexical items. In other words, and contrary to what has been and still largely remain a common practice in much lexicographic work on signed languages, we believe that adequate dictionaries need to be based on extensive corpora of signed texts of different genres, along the lines pointed out by Russo (2005). In addition, in our view, it would be very useful to create and analyze not only transcriptions of signed data (which reflect the equivalent of the “oral” modality of spoken language use), but also corpora of texts *conceived and expressed directly in a written form*, as exemplified above.

We have found that many insights on the structure of LIS lexicon and grammar can be gained by reflecting on the structure of texts, on how the individual components of a text need to be segmented and are at the same time interrelated to express meanings. Comparing the individual units identified in text corpora, and examining how their form changes or remains unaltered, depending upon the grammatical and discourse context, is a powerful theoretical-methodological tool for identifying “citation forms” that may eventually be used for creating dictionaries based on actual usage, as suggested by Russo (2005).

At the same time, it is quite obvious that the actual capability of a written representation system (regardless of its use as a writing or transcription tool) must be tested on *both* individual signs and textual units. Thus in principle, as well as for practical purposes, the problem of representing corpora of individual signs (as when building dictionaries) cannot and, in our view, should never be separated from the problem of representing corpora of signed texts.

We also believe that, in order to be appropriately addressed, the issue of representing signed languages requires a profound metalinguistic awareness of “writing” as distinguished from “transcribing”. This distinction is often taken for granted in spoken language research, but is rarely made clear in research on signed languages. We strongly believe that a thorough awareness of this distinction is quite crucial when dealing with four-dimensional languages that have not spontaneously evolved a written form, such as our language, LIS.

6. References

Bergman, B., Boyes Braem, P., Hanke, T., & Pizzuto (guest eds.) (2001). *Sign Language and Linguistics*,

(*Special Issue on Sign Transcription and Database Storage of Sign Information*), 4: 1/2.
Breton, R. (2003). *Atlas des langues du monde, Une pluralité fragile*. Paris: Éditions Autrement.
Cecco, S. (2001). SignWriting: Un sistema di scrittura per le lingue dei segni. Verso un'applicazione alla Lingua dei Segni Italiana. Unpublished Dissertation, University of Trieste, Scuola Superiore per Interpreti e Traduttori.
Cuxac, C. (2000). *La langue des signes française (LSF); les voies de l'iconicité.* Faits de Langues 15/16, Paris: Ophrys.
Emmorey, K. (ed.) (2003). *Classifier constructions in sign languages*. Mahwah, NJ: Erlbaum.
Fabbretti, D., & Pizzuto, E. (2000). Dalle mani alla carta: Aspetti teorici e metodologici della notazione della Lingua Italiana dei Segni. *Rassegna di Psicologia*, 2, vol. XVII, 101-122.
Garcia, B. & Dalle, J. (2005). Le projet français "LS-script" pour une formalization graphique des langues des signes, et l'experience SW en France. Seminar given at the *ISTC-CNR & ISSR*, *2005-2006 Seminars' Cycle,* Rome, October 27, 2005.
Newkirk, D. (1989). *SignFont Handbook*. Bellevue, WA: Edmark Corporation.
Ochs, E. (1979). Transcription as a theory. In E. Ochs, B. B. Schiefflin (eds.), *Development pragmatic*s, New York: Academic Press, 43-72.
Ong, W.J. (1982). *Orality and literacy. The technologizing of the world*. London: Methuen.
Pennacchi, B., Pizzuto, E., & Rossini, P. (2001). La LIS: trascrizione e scrittura. Seminar given at the Institute of Psychology, CNR, Rome - *2000-2001 Seminars' cycle of the Division of Neuropsycology and Deafness & The State Institute of the Deaf*, April 2001.
Pietrandrea, P. (2000). Sistemi di analisi e rappresentazione delle lingue dei segni. Seminar given at the Institute of Psychology, CNR, Rome - *1999-2000 Seminars' cycle of the Division of Neuropsycology and Deafness & The State Institute of the Deaf*, February 2000.
Pietrandrea, P. (1997). I dizionari della LIS: analisi quantitative e qualitative. In M.C. Caselli & Serena Corazza (eds.), *LIS – Studi, esperienze e ricerche sulla Lingua dei Segni in Italia*. Pisa: Edizioni del Cerro, 255-259.
Pizzuto, E. & Cuxac, C. (2004). *Language, its formal properties and cognition: What can be learned from signed languages*. Joint Italian/French CNR/CNRS Project (ISTC-CNR & Université Paris 8 / CNRS Unit UMR-7023 - "Structures Formelles du Langage")
Pizzuto, E, Pennacchi, B. e Rossini, P. (to appear), Representing signs in written form: issues to be addressed. *Proceedings of the International Congress on Aprendizaje de las Lenguas Europeas en personas Sordas y Sordiciegas en el siglo XXI - APREL XXI, European Union, European Year of Languages 2000,* University of Barcelona, 30 November - 1 December 2001.
Pizzuto, E., Pennacchi, B., Rossini, P., Sutton, V. , Pietrandrea, P., Russo, T. (2000). An Issues Raising Approach applied to different written representations of a single LIS utterance. Paper presented at the Workshop "Cross-linguistic investigations of signed languages: can similarities and differences be detected without appropriate tools for representing and analyzing signed texts?". *7th International Congress on Theoretical Issues in Sign Language Research*. Amsterdam, July 23-27, 2000.
Pizzuto, E., & Pietrandrea, P. (2001). The notation of signed texts: open questions and indications for further research, *Sign Language and Linguistics*, (*Special Issue on Sign Transcription and Database Storage of Sign Information*), 4: 1/2, pp. 29-43.
Pizzuto, E., Rossini, P., Russo, T., Wilkinson, E. (2005). Formazione di parole visivo-gestuali e classi grammaticali nella Lingua dei Segni Italiana (LIS): dati disponibili e questioni aperte. In M. Grossmann & A. M. Thornton (eds.), *La formazione delle parole.* Roma: Bulzoni, 443-463.
Prillwitz, S., Leven, R., Zienert, H., Hanke, T., Henning, J., *et al*. (1989). *Hamburg Notation System for Sign Languages. An introductory guide*. HamNoSys version 2.0. .Hamburg: Signum Press.
Radutzky, E. (1997). La configurazione, In M. C. Caselli & Serena Corazza (eds.). *LIS – Studi, esperienze e ricerche sulla Lingua dei Segni in Italia*. Pisa: Edizioni del Cerro, 87-111.
Russo, T. (2004). *La mappa poggiata sull'isola. Iconicità e metafora nelle lingue dei segni e nele lingue vocali*. Rende: Università della Calabria.
Russo, T. (2005). Un lessico di frequenza della LIS. In T. De Mauro & I Chiari (eds.), *Parole e numeri*. Rome: Aracne, 277-290.
Sutton, V. (1999). *Lessons in SignWriting. Textbook & workbook* (2nd edition, 1st edition 1995). La Jolla, CA: Deaf Action Commitee for Sign Writing.
Volterra, V. (ed.). (2004). *La Lingua dei Segni Italiana*. Bologna: Il Mulino (first edition 1987).
Wilcox, S. & Wilcox, P. (1995). The gestural expression of modals in American Sign Language. In S. Fleischman & J. Bybee (eds.), *Mood and modality*. Amsterdam: John Benjamins.

7. Acknowledgements

We must thank Elena Pizzuto, from the depth of our hearts, for all the energetic support she has given us and for all the patience she has employed in trying to follow and understand our discussions. Our thanks go to her also for the great help she has given us in dealing with the unusual (for us) "art" of selecting and describing in a research paper (in a foreign language, too!) the many essential points that emerged during these meetings.

We wish to thank also Juliette Dalle, Brigitte Garcia, Nicolas Médin and Paolo Rossini, and the anonymous LREC reviewers, for their many remarks that helped us considerably during the preparation of this paper.

We gratefully acknowledge partial financial support from the Istituto Statale Sordi di Roma (ISSR) and the Association "Progetti Felicità" Famiglia Rana (Joint ISTC-CNR/ISSR project "Writing LIS & SignWriting"), and from CNR and CNRS (French- Italian Project "Language, its formal properties and cognition: what can be learned from signed languages" - ISTC-CNR, Rome & UMR 7023 CNRS, Paris).

High level models for sign language analysis by a vision system

Patrice DALLE
IRIT – Paul Sabatier University
118 route de Narbonne – 31062 Toulouse cedex 9 - France
E-mail: dalle@irit.fr

Abstract

Sign language processing is often performed by processing each individual sign and most of existing sign language learning systems focus on lexical level. Such approaches rely on an exhaustive description of the signs and do not take in account the spatial structure of the sentence. We present a high level model of sign language that uses the construction of the signing space as a representation of both (part of) the meaning and the realization of a sentence. We propose a computational model of this construction and explain how it can be attached to a sign language grammar model to help analysis of sign language utterances and to link lexical level to higher levels. We describe the architecture of an image analysis system that performs sign language analysis by means of a prediction/verification approach. A graphical representation can be used to explain sentence construction.

1. INTRODUCTION

As other languages, sign language relies upon several grammatical levels, namely lexical, syntactical, and semantical levels. However, most of the existing researches focus on the lexical level and moreover, on standard signs i.e, the ones that can be found in dictionaries, and not on iconic utterances (classifier or “proforms”, transferts structures, …). On the other hand, iconic structures are widely used in spontaneous sign language so it seems appropriate to take them in account in automatic sign language processing systems.

The meaning of a sign language production can be recovered by considering the construction of the signing space (Cuxac 1999) (Cuxac, 2000): during this production, the signer uses this space to position the entities that are evoked in the sentence and to materialize their semantic relationships, so that the resulting construction can be considered as a representation of the meaning of the discourse.

We propose a computational representation of this organization, and describe how this representation can be used to help automatic interpretation of sign language by an image processing system, and how graphical representation can help sign language understanding and learning.

Most of previous works on sign language analysis focused on isolated sign translation by means of a finite set of parameters and values, from the Liddel and Johnson phonological description (Vogler 1998) or the Stokoe description system (Ouhyoung 1998). Datagloves are often used as input devices. Some works focus on increasing the recognition rate by using some additional knowledge on the signed sentence structure: statistics on consecutive pairs of signs (stochastic grammars) (Hienz 1999) or (Ouhyoung 1996), constraints on the structure of the sentence (Pentland 1995). Nevertheless these approaches do not take in account the spatial structure of the signed sentence. The resulting systems are only able to deal with sentences considered as a simple succession of isolated signs, eventually coarticulated. More complex aspects of sign language such as sign space utilization or classifiers have not been studied yet in vision-based sign language analysis, but some issues where brought out in recent works on sign language generation (Bossard 2003) (Huenerfauth 2004).

Our approach focuses on the fact that introducing knowledge about sign language syntax and grammar will allow a vision system to achieve the image analysis of the sequence and, thus, avoid us to systematically use complex reconstruction of gestures. Instead of direct sign recognition, we make much of identifying the structure of the sentence in terms of entities and relationships, which may be sufficient in a reduced-context application. This allows us to use a general model of sign language grammar and syntax. Hence, starting from an high level hypothesis about what is going to be said in the sign language sentence, this model let us compute a set of low level visual events that have to occur in order to validate the hypothesis. While verifying the fact that something has happened is simpler than detecting it, our approach permits the use of rather simple image processing mechanisms in the verification phase and reserves explicit reconstruction of gestures for the cases where prediction becomes impossible.

2. OVERVIEW OF OUR APPROACH

In order to analyse FSL utterances using a single video camera and simple image processing, we need to integrate a fair amount of knowledge (i) about FSL grammar and syntax for prediction and consistency checking of the interpretation but also (ii) about image processing for querying the low-level verification module.

The system integrates this knowledge in a multi-level architecture that is divided in three main subsystems:
1. The first subsystem consists in a representation of the interpretation of the discourse through a modeling of the signing space. During processing, the coherence of signing space instantiation is controlled by a set of possible behaviors resulting from the structure of the language and from a semantic modeling of the entities in the discourse.

2. The second subsystem is a knowledge representation system based on description logic formalism. The base contains some knowledge about FSL grammar and syntax that makes it able to describe high level events that occurred in signing space in terms of low level sequences of events on body components.

3. The last subsystem performs image processing; it integrates knowledge about the features it must analyze so as to choose the appropriate measurement on the data for the verification process.

Next sections describe the main aspects of the linguistic model and the verification process.

3. MODELING THE SIGNING SPACE

3.1 SIGNING SPACE MODEL

In the FSL, entities are evoked through signs and located in the signing space so that their relative positions will correspond to spatial relationships between those entities in the real world. Temporal relationships are evoked through entities that are located on "time lines". Binary actions are evoked through directional verbs and more complex ones by grammatical structures called "transfers" (Cuxac 1999). The different kinds of entities depend on the kinds of relationships in which each entity may be involved: *dates* can be involved in temporal relationships, *places* in spatial relationships; *animates* can perform an action or be located relative to another entity, *actions* can be referenced as a moment in time or as one of the protagonists of an action. The specificities of the FSL grammar require to consider some additional kind of entities: one needs to make a distinction between entities that whenever involved in a complex action are evoked by the signer taking their role (*persons*1) and the entities that cannot be evoked this way (*objects*). Finally, due to the temporal ordering of the signs, one needs to take in account the case of actions that are evoked before one of their protagonists; the type of this entity is *implicit*.

3.2 SIGNING SPACE REPRESENTATION

The symbolic representation of the signing space consists of a volume surrounding the signer, regularly divided into Sites. Each location may contain a single Entity, each Entity having a Referent. A Referent is a semantic notion that can be found in the discourse. Once it has been placed in the signing space, it becomes an Entity and has a role in the sentence. Hence, building a representation of a sign language sentence consists in creating a set of Entities in the SigningSpace. A graphical representation of the signing space can be built to explain how FSL uses the space as seen in figure 1.

The meaning contained in this signing space construction is represented in terms of Entities whose Referents can have successively different function(s) during the construction of the sentence (locative, agent, actions, …). A set of rules maintains the consistency of the representation by verifying that sufficient coherent information has been provided when one needs to create a new entity in the signing space. The global architecture of the model can be represented in UML notation standard.

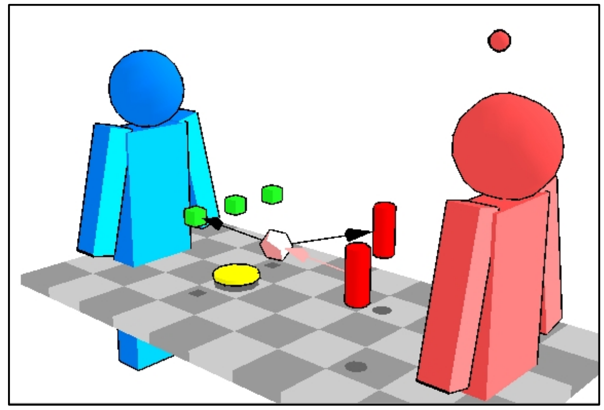


Figure 1: Signing space representation

4. A MODEL TO CONSTRUCT THE SIGNING SPACE

4.1 RULES OF THE SIGNING SPACE CONSTRUCTION

We need rules of FSL grammar to describe the signing space construction. As modifying the signing space only consists in creating new entities, our model focuses on the gestures that are used to create those entities. Without lexical knowledge, it is not possible to make a distinction between entities that are neither dates nor actions. So that creating such an entity relies on a generic mechanism.

Creating an entity of a given type relies on the following mechanisms:

• Creating a generic entity: entities are created and localized in the signing space by signs that can be performed either directly in the desired location or localized on the signer's body for lexical reasons. In the second case, the production of the sign is followed by an explicit designation of the desired location.

• Creating a date: in our reduced context, dates are explicitly evoked by standard signs, performed in a neutral location (if front of signer's chest) and located simultaneously on one of the time lines.

• Creating an action: binary actions are evoked through directional verbs, which implies some gestures that explicitly connect two locations containing entities in the signing space. For complex actions, "great iconicity" structures such as those where the signer plays the role of one of the action's protagonist have to be used. We have not yet study such complex actions. .

The formalization of that grammar relies on the fact that each of those mechanisms can be described by a gesture sequence.

4.2 DESCRIBING THE CONSTRUCTION OF THE SIGNING SPACE

A modification in the signing space is defined by the kind of the entity that is created and its localization. The behavior model attaches to each kind of entity a gesture sequence that describes the state of the components involved and the way they are synchronized.

The computational representation of that grammar relies on a description logic formalism and uses the CLASSIC knowledge representation system (Brachman 1991). This system expresses the representation of FSL grammar as a set of hierarchically organized concepts. Concepts are structured objects, with roles (concepts of a given type) and associated with automatic inference mechanisms and user-defined propagation rules.

On the basis of the description logic formalism, describing the creation of an entity consists in defining a set of concepts with specific constraints on some of their roles:
1. The concept representing the creation of an entity is called ACTS (ACtion Transforming Signing space). It is described by a location, a temporal interval and a gesture sequence.
2. Gesture sequences consist in a list of component descriptions associated with constraints on the values of the component roles.
3. Additional knowledge propagation rules concern vertical information propagation from an ACTS description to gestures defined in the corresponding sequence (e.g. the localization of the hand must be the same as the one of the entity). Horizontal information propagation mechanisms are used between different gesture descriptions in the same sequence (e.g. both hands must have the same location). Finally gestures synchronization rules are based on Allen's algebra operators.
This formalization leads to a global representation of the FSL grammar as a concept hierarchy associated with additional propagation rules sets (figure 2)

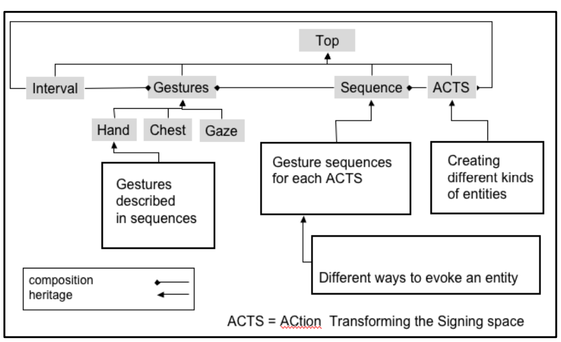


Figure 2 Concept hierarchy and inference mechanisms

For each kind of entity, there is a specialization of the ACT concept with a specific *GestureSequence*. This sequence can be derived depending on the different ways of creating an entity of that type. Gestures that can be found in *GestureSequences* are specializations of generic *Component* descriptions that include additional constraints on their roles.

5. IMAGE-BASED SIGN LANGUAGE ANALYSIS

The representation of the signing space can be linked to the meaning of the discourse by giving access to the relationships between entities that were evoked and referenced. On the other hand, the iconicity theory by (Cuxac 1999) provides a description of the grammar of the sign language in terms of gesture sequences that leads to creating a new entity in the signing space. As a result, this permits to link this representation to the gestures that were used to create the current signing space instantiation. Such a predictive model can be used for analysis of sign language sentences.

Using that model for sign language analysis leads to two classes of tools: (i) interactive tools intended for linguists to evaluate the model or for teachers to explain sign language, (ii) automatic analysis tools that can be used in many fields of application (linguistic analysis, automatic interpretation,).
An interactive tool has been developed in order to represent the construction of the signing space during the production of the utterance (fig. 3). This tool consists of a transcription software that allows to synchronously link the different steps of the construction of the signing space and the video sequence that is transcripted. This application was designed to evaluate the model with respect to several kinds of utterances and to determine how this model can be considered as a generic representation of sign language utterances.

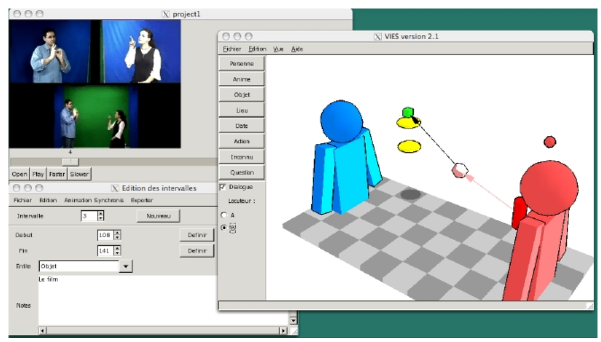


Figure 3 : interactive tool to build signing space

In the field of automatic analysis, using a single camera, it is not possible to build an exhaustive description of the gestures that are used. Therefore, automatic vision-based sign language analysis, the model of the signing space is used as a general representation of the structure of the sentence that simultaneously gives access to the meaning of the discourse.
The grammar of the sign language that can be attached to this construction allows the use of a prediction verification approach (Dalle 2005): from an hypothesis on the meaning of the discourse in terms of a signing space modification, it is possible to infer the gestures that were used to create the new entity in the signing space. Analyzing the utterance is then reduced to verify whenever the data corroborates this prediction or not.
Such an analysis can be performed without taking in account the lexicon, so that the gestures descriptions that can be used need to be less precise that the ones required for exhaustive sign recognition. This makes the analysis of low resolution images possible.

However, in a reduced context, the spatial structure of the sentence may be an interesting guideline to identify the signs as it can be done by only considering discriminative aspects of the signs. The behavior model infers a gesture sequence and asks the image processing module to verify it. The system describes each item of the gesture sequence in visual features. This reformulation is made in a qualitative way. For instance it does not need an exact knowledge about hand shape, but only to know whether it is changing or not. Then, each of these features can be verified using simple 2D clues. For instance, to test hand shape properties, we only have to consider simple 2D shape properties as area or bounding box; to test if the signer looks at the location of the entity, we measure the dissymmetry of the face from the chest axis. Without this prediction process, in a bottom-up analysis, we should have to extract and recognize arm movement or hand configuration and so, to use more complicated measures as 3D tracking trajectories, shape descriptors, gaze direction or 3D face orientation.

The three different elements of such automatic tool (signing space representation, grammatical model, low level image processing) have been evaluated separately. It has been shown that in a reduced context, the prediction/verification approach was relevant and allowed to use simple 2D image processing operators instead of complex gesture reconstruction algorithms to performs the identification of the different kinds of entities that where used in the utterance.

6. CONCLUSION

In conclusion, this model is our first formalization of spatio-temporal structure of the signing space. Its purpose is to help sign language image analysis.
The main interests of this approach are:

- the use of a qualitative description of the gestures that can be easily identified with simple and robust image processing techniques,
- the use of a prediction / verification approach where only significant events have to be identified and that avoid an exhaustive reconstruction of the gestures,
- the descriptions used in that model provide a strong guideline for the design of those operators.

Implementation of the model and tools we have built help linguists to evaluate their linguistic model of sign language and teachers to explain FSL structures..
Further works concern:

- The extension of the model to dialog situation, with shared entities,
- The implementation of more complex transformations as “transfer structures”.

Finally, signing space representation could be used for the specification of a graphical form of sign language.

7. REFERENCES

Bossard, B., Braffort, A., Jardino, M. (2003). Some issues in sign language processing. In A. Camurri and G. Volpe (Eds.), Lecture Notes in Artificial Intelligence : 5th International Workshop on Gesture-Based Communication in Human-Computer Interaction, (pp. 15-17) Genova, Italy

Brachman, R.J. and al. (1991). Living with classic: When and how to use a kl-one-like language. In J. Sowa, (Eds.), Principles of Semantic Networks: Explorations in the representation of knowledge, (pp.401–456). Morgan-Kaufmann, San Mateo, California.

Cuxac, C. (1999). French sign language: proposition of a structural explanation by iconicity. In A. Braffort, R. Gherbi, S. Gibet, J. Richardson and D. Teil, (Eds.), Lecture Notes in Artificial Intelligence : Procs 3rd Gesture Workshop’99 on Gesture and Sign-Language in Human Computer Interaction (pp. 165-184) Gif-sur-Yvette, France, Springer: Berlin

Cuxac, C. (2000). La langue des signes française (LSF). Les voies de l’iconicité, Faits de Langues n° 15-16, Ophrys, Paris.

Dalle P., Lenseigne B. (2005). Vision-based sign language processing using a predictive approach and linguistic knowledge. In: IAPR conference on Machine Vision Applications – MVA (pp. 510-513). Tsukuba Science City, Japon,. IAPR.

Dalle P., Lenseigne B. (2005). Modélisation de l’espace discursive pour l’analyse de la langue des signes. In TALS 2005, atelier de TALN 2005. [on line] [http://tals.limsi.fr/actes/s7.pdf].

Hienz, H., Bauer, B., Kraiss, K.F. Hmm-based continuous sign language recognition using stochastic grammars. In A. Braffort, R. Gherbi, S. Gibet, J. Richardson and D. Teil, (Eds.), Lecture Notes in Artificial Intelligence : Procs 3rd Gesture Workshop’99 on Gesture and Sign-Language in Human Computer Interaction (pp. 185-196) Gif-sur-Yvette, France, Springer: Berlin.

Huenerfauth, M. (2004). Spatial representation of classifier predicates for machine translation into american sign language. In Workshop on Representation and Processing of Sign Language, 4th Internationnal Conference on Language Ressources and Evaluation (LREC 2004), (pp. 24–31), Lisbon Portugal.

Ouhyoung M. Liang. R.H.(1996). A sign language recognition system using hidden markov model and context sensitive search. In Proceedings of the ACM Symposium on Virtual Reality Software and Technology, (pp . 59–66), Hongkong

Ouhyoung, M., Liang, R.H. (1998) A real-time continuous gesture recognition system for sign language. In 3rd International conference on automatic face and gesture recognition, (pp. 558–565), Nara, Japan,

Pentland, A. , Starner, T. (1995). Real-time american sign language recognitio from video using hidden markov models. Technical Report TR-375, M.I.T Media Laboratory Perceptual Computing Section.

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, (pp. 363–369), Mumbai, India.

Continuous Sign Language Recognition – Approaches from Speech Recognition and Available Data Resources

Morteza Zahedi, Philippe Dreuw, David Rybach, Thomas Deselaers, Jan Bungeroth and Hermann Ney

Lehrstuhl für Informatik 6 – Computer Science Department
RWTH Aachen University
D–52056 Aachen, Germany
`<surname>@cs.rwth-aachen.de`

Abstract

In this paper we describe our current work on automatic continuous sign language recognition. We present an automatic sign language recognition system that is based on a large vocabulary speech recognition system and adopts many of the approaches that are conventionally applied in the recognition of spoken language. Furthermore, we present a set of freely available databases that can be used for training, testing and performance evaluation of sign language recognition systems. First results on one of the databases are given, we show that the approaches from spoken language recognition are suitable, and we give directions for further research.

1. Introduction

The first generation of sign language recognition systems has employed special data acquisition tools like gloves or wearable cameras to obtain the features to recognize the gestures (Vogler and Metaxas, 1997; Starner et al., 1998; Bauer et al., 2000). Only few research groups use databases which have been recorded using normal stationary cameras (Bowden et al., 2004; Zahedi et al., 2005; Zieren and Kraiss, 2005). Nonetheless, most of the databases have been recorded in a highly restricted environment with constant lightning, homogeneous, non-changing background, and the signers are dressed in long-sleeve shirts. In such an environment motion- and skin-color detection is greatly simplified, resulting in a task that is only slightly more difficult than the tasks where data-gloves were used.

Some other databases have been created by linguistic research groups. These databases have not been produced with sign language recognition in mind; i.e., no suitable transcription is available. To use these data for the training or for performance evaluation in sign language recognition systems, the necessary transcriptions have to be created, which is a costly process and requires a lot of human work.

In this paper, we present different databases which have been prepared in different ways: (i) I6-Boston201 database: consists of 201 sentences of American sign language (ASL) and have been recorded in a controlled environment. The signs have been recorded by four standard stationary cameras. It is a subset of the database recorded by Boston University (Neidle et al., 2000). (ii) Phoenix database: has been recorded from the daily news “Tagesschau” of the German TV channel Phoenix. In this program an interpreter signs the news in German sign language simultaneously in the lower right corner of the TV screen. This database is transcribed in German sign language and German language. The movies are not recorded in a controlled environment, but instead the signer is shown in front of a strongly non-homogeneous, non-constant background. (iii) the ECHO database consists of three corpora: British sign language (BSL) (Woll et al., 2004), Swedish sign language (SSL) (Bergman and Mesch, 2004) and sign language of the Netherlands (NGT) (Crasborn et al., 2004), respectively. We have prepared the ECHO databases for sign language recognition by choosing some parts of the original corpora and creating the necessary annotations.

Our automatic sign language recognition system is derived from a large vocabulary automatic speech recognition system, because both, speech and sign language are sequences of the features over the time. In section 2, a short overview of the system is presented. We will introduce the databases in section 3 and finally preliminary results of the system and conclusion are shown in section 4 and 5.

2. System Overview

As mentioned above, our sign language recognition system is based on a large vocabulary speech recognition system (Kanthak et al., 2000b; Gollan et al., 2005). This allows us to easily use the techniques developed for speech recognition and transfer the insights from this domain into automatic sign language recognition. Common speech recognition systems are based on the Bayes’ decision rule,

$$\hat{w}_1^N = \arg\max_{w_1^N}\{Pr(w_1^N) \cdot Pr(x_1^T|w_1^N)\} \qquad (1)$$

where $\hat{w}_1^N$ is the sequence of words that is recognized, $Pr(w_1^N)$ is the language model, and $Pr(x_1^T|w_1^N)$ is the visual model (cp. acoustic model in speech recognition), x_1^T are the features for the time slots 1 to T.

Obviously, the features x_1^T have to be extracted in a different way than in speech recognition using techniques known from the image processing domain. To handle video files we use the FFmpeg library[1], which is able to handle a wide range of different video formats. Basic image processing methods are integrated into the system: thresholding, cropping, rotation, resizing to allow for a suitable selection of the region of interest in the videos; convolution, Sobel filters, smoothing to pre-process images. Furthermore, methods that were successfully used in gesture recognition were integrated: skin color models (Jones and Rehg, 1998) to locate faces and hands, motion detection

[1] `http://ffmpeg.sourceforge.net/index.php`

Figure 1: Sample frames from I6-Boston201.

by difference images (Dreuw et al., 2006), motion history images (Morrison and McKenna, 2004), geometric features (Rigoll et al., 1998), and spatial features (Bowden et al., 2004). In (Dreuw et al., 2006) a tracking algorithm using dynamic programming was introduced that considers the complete image sequence to find the best tracking-path with respect to a given criterion. This tracking can be used in the recognition process in the same way as time-alignment is used in speech recognition.

This framework allows for easy testing and development of new features for automatic sign language recognition. It is easily possible to reconfigure the system, to change parameters, to use different corpora and to change the feature extraction process. A description of the speech recognition system can be found in (Kanthak et al., 2000a).

3. Databases

In this section, three different sign language databases are presented. These databases are a starting point for performance evaluation in automatic sign language recognition. Where missing, we created the necessary annotation to be able to use them for automatic sign language recognition. All the data are freely available on the Internet.

3.1. I6-Boston201 Database

The National Center for Sign Language and Gesture Resources of the Boston University has published a database of ASL [2]. We have used 201 annotated videos of ASL sentences. Although this database was not recorded primarily for image processing and recognition research, we considered it as a starting point for a recognition corpus because the data are available to other research groups and can thus be a basis to compare different approaches. The database consists of videos from three signers: one male and two female signers. The signers are dressed differently.

The signing is captured simultaneously by four stationary standard cameras, three of them are black/white cameras and one is a color camera. All cameras have fixed positions. Two sample frames are shown in Figure 1.

Two black/white cameras, directed towards the signer's face, form a stereo pair that can be used to obtain three-dimensional data. Another camera is installed on the side of the signer.

The color camera is placed between the cameras of the stereo pair and is zoomed to capture only the face of the

[2] http://www.bu.edu/asllrp/ncslgr.html

	Training set		Evaluation set
	Training	Development	
Sentences	131	30	40
Running Words	695	172	216
Unique Words	103	65	79
Singletons	37	38	45

Table 1: Corpus statistics for I6-Boston201 database.

Figure 2: Left: whole screen image, right: close up to the interpreter.

signer. This camera can be used for facial expression analysis. The movies are recorded at 30 frames per second and the size of the frames is 312×242 pixels. We use the published video streams at the same frame rate but extract the upper center part of size 195×165 pixels. (Parts of the bottom of the frames show some information about the frame, and the left and right border of the frames are unused.)

To use these data for ASL sentence recognition, we separated the recordings into a training and evaluation set. To optimize the parameters of the system, the training set is further split into separate training and development parts. To optimize parameters in the training process, the system is trained by using the training set and evaluated using the development set. When parameter tuning is finished, the training data and development data are used to train one model using the optimized parameters. This model is then evaluated on the so-far unseen test set. This database is called I6-Boston201 in the following. Corpus statistics for this database are shown in Table 1 which include number of sentences, running words, unique words and singletons in the each part. Singletons are the words occurring only once in the set.

3.2. Phoenix Database

The German TV channel Phoenix broadcasts the daily "Tagesschau" news program in German and with a German sign language translation in the lower right corner of the screen. The whole screen and a close up of the interpreter are shown in Figure 2. We have recorded the complete "Tagesschau" for 104 days and currently a snapshot of the recordings consisting of the weather reports of 51 days is used. The sign language of these recordings is fully transcribed. These data are split into training, development, and test data and the complete corpus statistics of this database is given in Table 2. In total there are 11 different signers (1 male and 10 females).

The movies are in MPEG1 video format and in PAL res-

	Set		
	Training	Development	Evaluation
Sentences	421	79	56
Running Words	5890	500	389
Unique Words	643	168	139
Singletons	0	70	63

Table 2: Corpus statistics for Phoenix database.

Figure 3: Sample frames from ECHO databases.

olution (352×288). The database transcription has been created by a congenitally deaf using the ELAN software[3]. In addition to the pure transcription, information on the signers, start time and end time of the gestures and also boundaries of the sentences are available in the annotation files. Further information about annotation is available in (Bungeroth et al., 2006).

3.3. ECHO-Databases

The ECHO database[4] consists of three corpora in BSL, SSL and NGT. All three corpora include the videos from sign narrations of the same five fable stories, a small lexicon and interviews with the signers. In addition, there is sign language poetry in BSL and NGT. Figure 3 shows sample image frames. The corpora have been annotated linguistically and include sign language and spoken language transcription in English. In addition, SSL and NGT sections include Swedish and Dutch transcription, respectively.
Also these videos have been transcribed using the ELAN software and the transcription includes word and sentence boundaries for the sign language recognition.
To use the ECHO databases in the field of sign language recognition, we have chosen some parts of the five fable stories of the original database and have created a database for each of the subcorpora. We name these databases ECHO-BSL, ECHO-SSL, ECHO-NGT.
Although the data have been recorded in a completely controlled environment with constant background, it is currently very hard to use these three databases for sign language recognition: The number of singletons and the number of unique words are too high in relation to the total number of utterances. To reduce the data sparseness, we have decided to split the corpus into training and testing data only, i.e. for these corpora no development sets have been specified. Furthermore, the test set was selected to have no out-of-vocabulary words, i.e. each word in the test set is at least once in the respective training set. The training corpora consists of the sentences and also segmented words of them but evaluation contains only sentences.

[3]http://www.mpi.nl/tools/elan.html
[4]http://www.let.ru.nl/sign-lang/echo

	Training set	Evaluation set
Sentences	206	56
Running Words	2628	237
Unique Words	534	97
Singletons	343	57

Table 3: Corpus statistics for ECHO-BSL database.

	Training set	Evaluation set
Sentences	136	23
Running Words	2988	129
Unique Words	520	70
Singletons	280	44

Table 4: Corpus statistics for ECHO-SSL database.

3.3.1. ECHO-BSL

The ECHO-BSL database is signed by 2 signers (1 male and 1 female). Statistics of the corpus is shown in Table 3.

3.3.2. ECHO-SSL

The ECHO-SSL database is signed by a male signer. Statistics of the corpus is shown in Table 4.

3.3.3. ECHO-NGT

The ECHO-NGT database is signed by 3 signers (2 males and 1 female). Statistics of the corpus is shown in Table 5.

4. Preliminary Results

In this section we present some preliminary results on the I6-Boston201 corpus introduced in the previous section. For the experiments, the video frames were scaled down to the size of 32×32 pixels. The performance of the system is measured by the word error rate (WER) which is equal to the number of deletion, substitution and insertion of the words divided by the number of running words. The results on development and evaluation sets including the perplexity (PP) and WER of the system using different language models are shown in Table 6. The n-gram language models where the probability of a sentence is estimated from the conditional probabilities of each word given the $n-1$ preceding words are employed in the experiments. The n-gram language models are called zerogram, unigram, bigram and trigram where n is equal to 0, 1, 2 or 3, respectively.

	Training set	Evaluation set
Sentences	187	53
Running Words	2450	197
Unique Words	468	77
Singletons	268	40

Table 5: Corpus statistics for ECHO-NGT database.

Language Model	Development set		Evaluation set	
	PP	WER(%)	PP	WER(%)
Zerogram	105	75	105	65
Unigram	36	71	37	63
Bigram	8	68	9	57
Trigram	7	69	6	55

Table 6: Preliminary result of the system on I6-Boston201.

Currently, we are working with the other corpora and we are trying to find a suitable set of image features for good recognition results. Furthermore the parameters of the sign language recognition system have to be tuned towards the task at hand as the parameters that are used in speech recognition are not always suited for the recognition of sign language.

5. Conclusion

We have presented an overview of our current efforts in the recognition of sign language. In particular we have employed a large-vocabulary speech recognition system which was extended by basic image processing techniques and which is currently being extended with feature extraction methods for sign language recognition. Furthermore, we presented 5 different tasks which can be used to benchmark continuous sign language recognition systems. These databases are freely available and can thus be used by other research groups.

Acknowledgments

The authors would like to thank Georg Heigold and András Zolnay for fruitful discussion and lots of tips on how to use the speech recognition system. This work was partially funded by the Deutsche Forschungsgemeinschaft (DFG) under grand NE-572/6.

6. References

B. Bauer, H. Hienz, and K.F. Kraiss. 2000. Video-based continuous sign language recognition using statistical methods. In *Proceedings of the International Conference on Pattern Recognition*.

B. Bergman and J. Mesch, 2004. *ECHO Data Set for Swedish Sign Language (SSL)*. Department of Linguistics, University of Stockholm, http://www.let.ru.nl/sign-lang/echo.

R. Bowden, D. Windridge, T. Kabir, A. Zisserman, and M. Bardy. 2004. A linguaistic feature vector for the visual interpretation of sign language. In *Proceedings of ECCV 2004, the 8th European Conference on Computer Vision*.

J. Bungeroth, D. Stein, P. Dreuw, M. Zahedi, and H. Ney. 2006. A german sign language groups of the domain weather report. In *Proceedings of the 5th Iternational Conference on Language Resources and Evaluation*.

O. Crasborn, E. van der Kooij, A. Nonhebel, and W. Emmerik, 2004. *ECHO Data Set for Sign Language of the Netherlands (NGT)*. Department of Linguistics, Radboud University Nijmegen, http://www.let.ru.nl/sign-lang/echo.

P. Dreuw, T. Deselaers, D. Rybach, D. Keysers, and H. Ney. 2006. Tracking using dynamic programming for appearance-based sign language recognition. In *Proceedings of the 7th International Conference of Automatic Face and Gesture Recognition*.

C. Gollan, M. Bisani, S. Kanthak, R. Schlter, and H. Ney. 2005. Cross domain automatic transcription on the tc-star epps corpus. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 1, pages 825–828, Philadelphia, PA, March.

M. Jones and J. Rehg. 1998. Statistical color models with application to skin color detection. CRL 98/11, Compaq Cambridge Research Lab.

S. Kanthak, S. Molau, A. Sixtus, R. Schlüter, and H.Ney. 2000a. The rwth large vocabulary speech recognition system for spontaneous speech. In *Proceedings of the Konvens 2000*.

S. Kanthak, A. Sixtus, S. Molau, R. Schlter, and H. Ney, 2000b. *Fast Search for Large Vocabulary Speech Recognition*.

K. Morrison and S.J. McKenna. 2004. An experimental comparison of trajectory-based and history-based representation for gesture recognition. In *Proceedings of the International Gesture Workshop*.

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

G. Rigoll, A. Kosmala, and S. Eickeler. 1998. High performance real-time gesture recognition using hidden markov models. In *Proceedings of Iternational Gesture Workshop*.

T. Starner, J. Weaver, and A. Pentland. 1998. Real-time american sign language recognition using desk and wearable computer based video. *Transaction of Pattern Analysis and Machine Intelligence*, 20(2):1371–1375.

C. Vogler and D. Metaxas. 1997. Adapting hidden markov models for asl recognition by using three-dimentional computer vision methods. In *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*.

B. Woll, Sutton-Spence, and D. Waters, 2004. *ECHO Data Set for British Sign Language (BSL)*. Department of Language and Communication Science, City University (LONDON), http://www.let.ru.nl/sign-lang/echo.

M. Zahedi, D. Keysers, T. Deselaers, and H. Ney. 2005. Combination of tangent distance and an image distortion model for appearance-based sign language recognition. In *Proceedings of DAGM 2005, 27th Annual meeting of the German Association for Pattern Recognition*.

J. Zieren and K.F. Kraiss. 2005. Robust person-independent visual sign language recognition. In *Proceedings of the 2nd Iberian Conference on Pattern Recognition and Image Analysis*.

A Software System for Automatic Signed Italian Recognition

Ignazio Infantino[1], Riccardo Rizzo[1], and Salvatore Gaglio[1,2]

[1]Istituto di Calcolo e Reti ad Alte Prestazioni, Consiglio Nazionale delle Ricerche, sede di Palermo

Edif. 11, Viale delle Scienze, 90128, Palermo, Italy

[2]Dipartimento di Ingegneria Informatica, Università degli Studi di Palermo,

Edif. 6, Viale delle Scienze, 90128, Palermo, Italy

E-mail: infantino@pa.icar.cnr.it, ricrizzo@pa.icar.cnr.it, gaglio@unipa..it

Abstract

The paper shows a system for automatic recognition of Signed Italian sentences. The proposed system is based on a multi-level architecture that models and manages the knowledge involved in the recognition process in a simple and robust way, integrating a common sense engine in order to deal with sentences in their context. In this architecture, the higher abstraction level introduces a semantic control and an analysis of the correctness of a sentence given a sequence of previously recognized signs. Experimentations are presented using a set of signs from the Italian Sign Language (LIS) and a sentence template useful for domotic applications, and show a high recognition rate that encourages to investigate on larger set of sign and more general contexts.

1. INTRODUCTION

In the recent years various approaches dealing with different sign languages have been proposed: many of the works are based on American Sign Language (ASL), but there are some works on different languages as the Chinese (Ma et al., 2000; Wang et al., 2000), German (Bauer et al.,2000), Netherlands (Grobel & Assan, 1997), Taiwanese Language (Liang & Ouhyoung, 1998), and English language (Sweeney & Downton, 1997).

From the computer engineering point of view, the works about Sign Language Recognition has been focused on two main tasks: the single sign recognition and the continuous sign recognition. The first one deals with the reconstruction and classification of a determinate configuration of arms, hands, head, and body from a static snapshot, or a movement that represents a single word. A large part of the works on this problem uses neural networks techniques with various approaches and configurations (Hamilton & Micheli-Tzanakou, 1994; Kim et al., 1996; Yang et al., 2002; Wilson & Anspach, 1993). Other works are based on the recognition of a hand posture from a static image using the appearance approach or model based approach (Cui & Weng, 2000; Triesh & von der Malsburg, 1996; Waldrom & Kim, 1995). The second research issue, i.e. the analysis of the sequence of signs, concerns the analysis of complex movements as continuous signals in space and time. Many of the works (see for example Vogler & Metaxas, 2001) use statistical approach based on Hidden Markov Models (HMMs), exploiting the assumption of the whole movement of a sign is decomposable in simpler components: chereme, i.e. the analogous of the phoneme in the sign language (Sweeney & Downton, 1997). Both single and continuous sign recognition need to detect and tracking spatial position (2D or 3D) of hands, arms, head of the signer. Often, this problem is easy resolved using marked gloves or gloves with special sensors. For example in Vogler & Metaxas, 2001, a set of magnetic markers were used in order to acquire the movement of hand and fingers. Recently, the non-manual gestures components are treated as a part of the sign language and processed together with the manual signs (see for example Grobel & Assan, 1997).

In this paper, we propose a inexpensive system that does not require special equipments or acquisition apparatus, but assure good performance using a sort of semantic context in a continuous sign language recognition task. The set of the admissible words, for a given detected sign, is disambiguated using the local context of the sentence. In order to perform that, we integrate a common sense engine in the process.

We explain the various functionalities implemented as a generic framework based on a cognitive architecture that allows to model and manage the knowledge of the recognition process in its wholeness. At the top of this architecture is the linguistic level that introduces the semantic context and allows the analysis of the correctness of a sentence given a sequence of recognized signs. Due to the predefined sentence templates that are used, the recognition is limited to a standardized gesture corresponding to words from the Italian Sign Language.

The paper is organized in five sections. After the introduction, the Sign Language features are highlighted to understand the implementation problems to approach. Section 2 deals with the various modules that compose the proposed system. Section 4 deals with a complete example of a sentence recognition. Finally, conclusions and discussion are reported.

2. FEATURES OF SIGN LANGUAGES

The interest in sign languages born with the Stokoe's work (Stokoe, 1978) that dealt with American Sign Language (ASL). The author discovered in ASL an organization similar to the common language, where a combination of simple sounds (phoneme) is used to create a very large number of words: in ASL a combination of simple gestures are used to generate a large number of signs with very different meanings. According with the analysis proposed by Stokoe a sign can be decomposed in three

parameters: the place in the space here the hands execute the sign (TAB); the configuration of the hands when they executes the sign (DEZ); the movement used to execute the sign (SIG). Another important parameter highlighted by Stokoe is the orientation of the palm, because some signs has the same DEZ, TAB and SIG but differs only for the palm orientation. It is to mention that in the set of signs of the ASL it is possible to identify the so-called Minimal Pairs: a couple of signs that differs only for a minimal variation of a single parameter. These non-manual components are very important in all the sign languages studied, because they convey an extra-linguistic information that is essential for the single sign recognition and also for continuous sign recognition.
In Italian Sign Language (LIS) the communication between the deaf and the other people the labial reading is frequently used since in the Italian language there is a very tight connection between the written and spoken language. The labial component is used in sign recognition when two sign are similar (Volterra, 1997).

3. SYSTEM OVERVIEW

The problem of sign translation must necessarily consider several aspects, known in literature as Recognition of Continuous Sign Language: the sign can begin and end in any instant of an observed sequence, since a temporal restriction in the execution of a sign does not exist; different signs have a variable duration, or the same sign can be executed with different duration; the transition from a sign to another is not exactly identified; a sign depends from the previous sign and the next sign (*coarticulation problem*); the begin and end of the single phrase is difficult to identify, the number of the signs in a phrase is not fixed.
All the cited problems make the recognition of continuous sign language a complex problem and a global solution is difficult to find. The aim of this work is to provide an architecture and a methodology to find a solution in the case of a single sentence. First of all we defined a subset of LIS sign that could be used in a home automation context. Such set is composed of verbs showing the key action to be performed, substantives, adjectives and time adverbs The system is based on a extensible words vocabulary, large enough to create a set of commands to give orders to a domestic robot or at a domotic automation system.
The proposed framework works in three main steps: sensorial input is processed in order to obtain a features vector by standard image processing algorithms; a Self Organizing Map (SOM neural network) is used to classified the single sign and to provide a list of probable meanings; a common sense engine finally choose the right meaning depending of the whole sentence context. An overview of the whole system architecture is depicted in figure 1. The system acquires a video sequence from a single video camera placed in frontal position with respect to the Signer.
The recognition is based on the space position of the signer's hands with respect to her head (Bowden et al., 2004; Charanyapan & Marble, 1992; Vogler & Metaxas, 2001). The main functionalities implemented are: the extraction and coding of the head and hands movement of the signer; the segmentation of the video input in single signs; the recognition of a single sign; the reconstruction of the whole sentence; the semantic analysis of the reconstructed sentence.
The incoming video is processed by the Movement detection Subsystem (MDS) that is the main component of the Pre-Processing Module (PM). This subsystem segments the video into several parts each of them representing a single sign, and generates a vector v. This vector is the input of the Sign Recognition subsystem that is in the Pattern Classification Module (PCM). This subsystem provides a list of the possible meanings for a single detected sign.
The last subsystem of the Pattern Classification Module collects the lists of meanings and reconstructs a set of sentences that are sent to the Sentence Recognition subsystem in Reasoning Module. This subsystem uses a common sense inferential software to check the meaning of the sentence. In the following subsections the implementation details of each single area are reported.

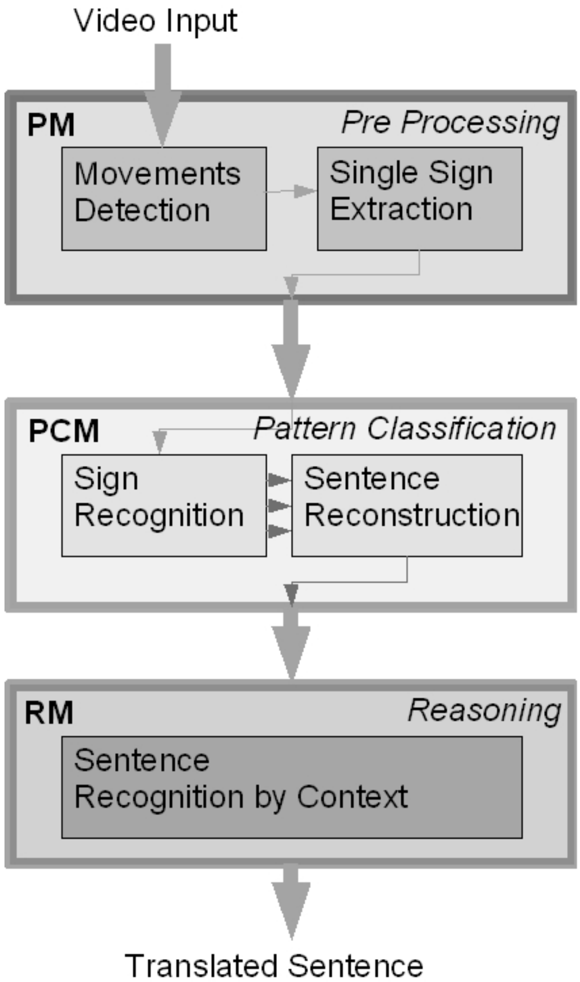


Figure 1: An overview of the proposed system.

The system is composed by the following processing blocks: Pre-Processing Module (PM), the Pattern Classification Module (PCM), the Reasoning Module (RM).

SIGN DETECTION ON IMAGE

The first step of the process is the pre-processing phase by the module PM. It aims to identify all the relevant image features (Bauer et al., 1990): we separated the skin areas of the image by using a colour based segmentation (see figure 2). The localization of head and hands is obtained by an algorithm based on detection of connected components. Only the three more extended regions are considered using the following criteria: head has greater area; hands are the others two regions if their extension is up to 10% of head area: in this way, the system is able to ignore non relevant regions; left and right hands are distinguished by their position on image. Each region is described by the following parameters: coordinates of the centroid ; width and height max ; width and height at the coordinates of the centroid; region area. The system follows the variation of these parameters, even if the occlusions happen.

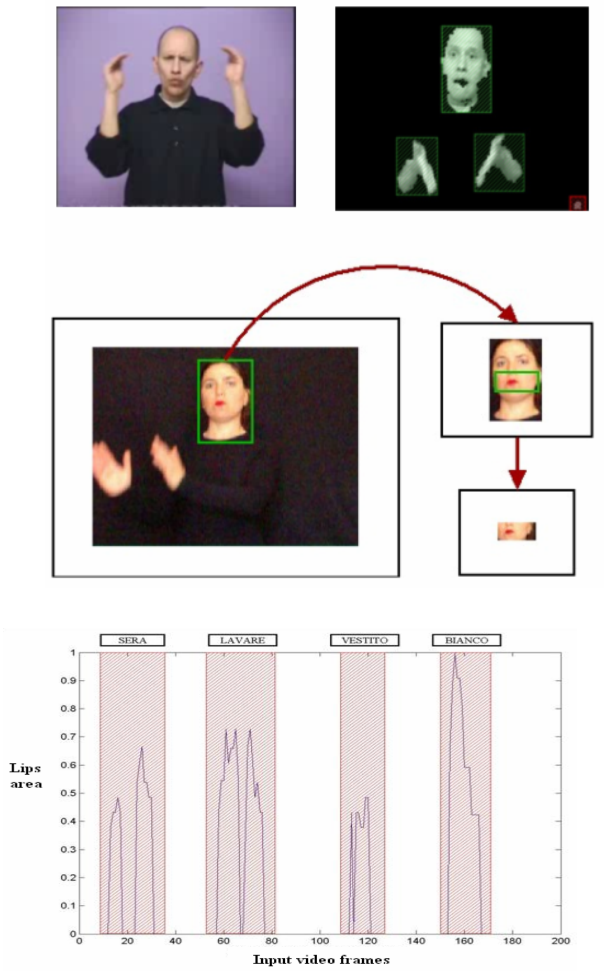


Figure 2: Pre-processing steps and results.

The pre-processing phase detect skin areas on the image and extract hand movements (first row). The motion of lips (middle row) is used to insulate single signs in a sentence (bottom row).

The data are collected for each frame of the sequence, and the whole movement corresponding to a sign is coded by a vector of real numbers. From the analysis of a huge number of sign videos the maximum sign duration observed is about 1300 ms. According to a given frame rate of 15 frames per second, a typical value even for a low-cost system, we choose n=10 a value that allows to capture the movements details. Having a fixed value of n allows obtaining a feature vector of 40 elements.

As said in section 2, the Italian signers often use lips movements during sign execution, especially when communicate with other people, and they don't feel as limiting to do it when use the system. This can be successfully exploited to solve the sentence segmentation problem. The pauses between lips movements reproducing the word of corresponding sign allow us to segment the sentence in fast and robust way (see figure2). This segmentation can be refined using an empirical consideration: usually an absence of signal can be considered a pause if it is about 500 ms. Considering frames rate of 15 fps this allows us to ignore signal interruptions that are shorter than 8 frames.

SIGN REPRESENTATION

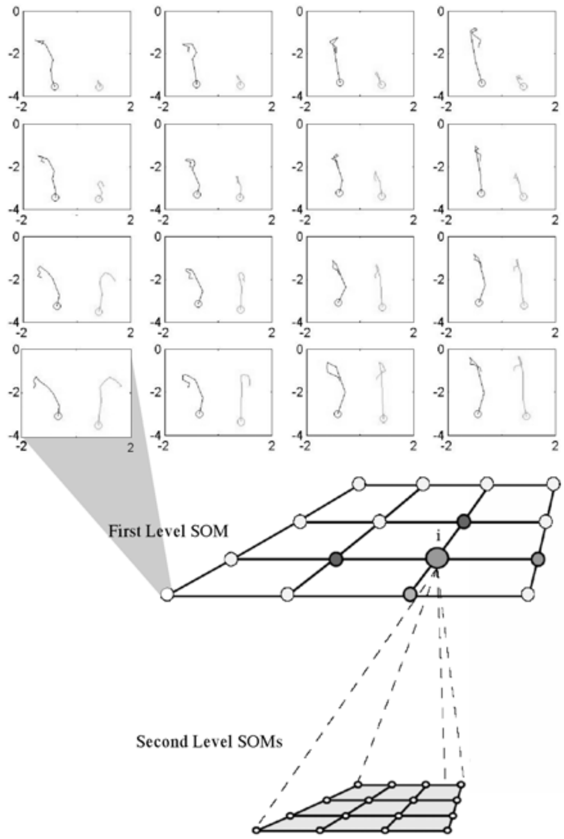
The aim of the Pattern Classification Module (PCM) is to work as a bridge connecting the perception to the symbolic processing, and to label the incoming sign representation. Using the signs representation as a vector it is possible to transform the comparison operation in a metrics measurement, and clustering allows to easily labelling the incoming pattern.

In our system the sentence recognition is made by the linguistic area, so the meaning of the sign is decided at the end of the processing chain, when the sentence is analyzed as a whole. To obtain this it is necessary to have a ranked list of possible meanings for each sign instead of a single answer (label) from the sign recognition block. This ranked list constitutes a search space for the Reasoning Module, and it is obtained exploring the representation space to find, inside the training set, similar signs. A topological map, obtained using a SOM neural network (Kohonen, 1997), supports the exploration of the representation space.

Using SOM, if an arbitrary pattern is mapped onto a unit, all points in a neighbourhood of it are mapped either to itself or to one of the units in the neighbourhood. This property is highlighted in the upper section of figure 3: the input patterns of our SOM are vectors that represent hands movements, and the weight vectors approximating these trajectories are visualized. Notice that the movements that involve only the right hand are mapped on the upper left corner of the map.

A SOM clustering system is usually obtained training the SOM with a set of patterns, then labelling the SOM units. This labelling can be done using the labels associated with the training patterns (in our case the meaning of the sign) using a voting criteria. Usually patterns that are neighbourhood in space have the same label or different label that can have a semantic relationship. If this property exists it is transmitted to the topological mapping. In our application due to the presence of the so-called *minimal pairs* (a sort of false friends for sign language), similar patterns can have very different label. Moreover for a single sign we can have few examples because for the user is a waste of time to repeat many and many times the same sign.

Many training of SOM with different topologies and different learning parameters were performed and a good compromise was obtained using a 4 X 4 topology. This allows to have a gross grain separation of the patterns, then, in order to obtain a finer separation of the signs inside the clusters, another layer of SOM networks was used (Miikkulainen, 1990). In second layer there is a SOM network for each unit of the first layer; i.e. the second layer is made by 16 SOM network of the same 4 by 4 topology (see figure 3). The hierarchical SOM classifier uses a total of 272 neurons, but is more efficient of a single layer SOM network. The SOM multilayer classifier associates a ranked list of labels to the input pattern (the single sign detected); this ranked list came from the labels



associated to the training patterns.

Figure 3. SOM multilayer structure.

The bottom part of the figure shows the SOM multilayer structure. The second level SOM will be trained using the training patterns of the gray units of the first level. In the upper section the weights of the neurons are visualized as the corresponding movement. In each square there is the movement of the hands, for example it can be noticed that signs with only right hand are in the upper-left corner.

COMMON SENSE CONTEXT ANALYSIS

The Reasoning Module Area is responsible for the syntactic control of the sequence of the signs, and generates the most probable sentence referring a given context. This module is based on OpenCyc[1], the open source version of the Cyc technology that implements a complete general knowledge base and commonsense reasoning engine. OpenCyc has been used as the basis of a wide variety of intelligent applications and expert systems. The principal functionalities of the module are: verification of the semantic correctness of a complete sentence; search of a single error in the sentence and suggestion of possible substitutes; correction of the error by evaluation of the most probable sentence using all the possible substitutions. Moreover the system is able to suggest next sign (or a set of possible signs) if an incomplete sentence is given. In the case of man-machine interaction, i.e. signer and recognition system are involved in a simple structured discussion (command/ request/ question+ answer+ …), the general context is defined but every recognized sentence generates a current context useful for process the next sentence. The context is introduced in Cyc defining a *microtheory*, i.e. a constant denoting assertions which are grouped together because they share a set of assumptions. It is accessible by querying to the Cyc Server about the truth of a sentence (formula), which may or may not contain an undefined sign (free variable). If the formula contains variables, then Cyc server supplies bindings to those variables which make the formula true (correct sentence); otherwise, it simply answers whether the formula is true. In the following experimental part examples of query are reported explaining the various functionalities of semantic control.

4. EXPERIMENTATIONS

The system is mainly written in Java code: it recalls the Preprocessing Module implemented using Matlab, that also includes the video acquisition capabilities; a specific Java class of the Reasoning Module manages the queries and the answers to/from the OpenCyc engine. The SOM multi-layer classifier was implemented using the Matlab SOM Toolbox.

The video sequences used for the training and the recognition has been acquired with a digital camera using a resolution of 320*240 at 15 fps (Pentax Optio 330GS). Moreover, we have mainly used some low resolution videos from a free database of sign ("Dizionario Italiano dei Segni", DIZLIS[2]). The implemented system has optimal performance if the signer is in front of the camera and the background is uniformed colored. Figure 4 shows the implemented graphical user interface that allows to manage all the computation.

The performances of the Pattern Classification Module are due to the Sign Recognition subsystem and the Sentence segmentation. The Sign Recognition subsystem is based on the SOM classifier and performances are determined by the training set that constitutes the vocabulary of the system. The larger vocabulary tested is made by 40 signs and there were added few minimal pairs. Each sign was repeated 4 times in order to obtain a set of 160 video fragments. The data extraction procedure generates a matrix of 160 rows and 40 columns. The training set is small compared with the dimension of the input space but to have more video samples is a problem because the user should repeat a single sign a lot of times. To artificially add more vectors to the training set we replicated the same representing vectors 4 times adding Gaussian noise with zero means and 0.1 variance. Using this method a matrix of 640 rows and 40 columns was obtained. Each row was labelled with a reference to the original video segment.

We submitted to the system 80 videos that where not part of the training set of the system and take the first five

[1] Cycorp, Inc. OpenCyc, http://www.cyc.com.

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

labels obtained from the Pattern Classification Module: for 50 videos the correct label was the first one, for the remaining videos the second label was checked and it was found correct in 17 cases, and so on. For two videos the correct label was not on the first five choices and these videos are considered not classified.

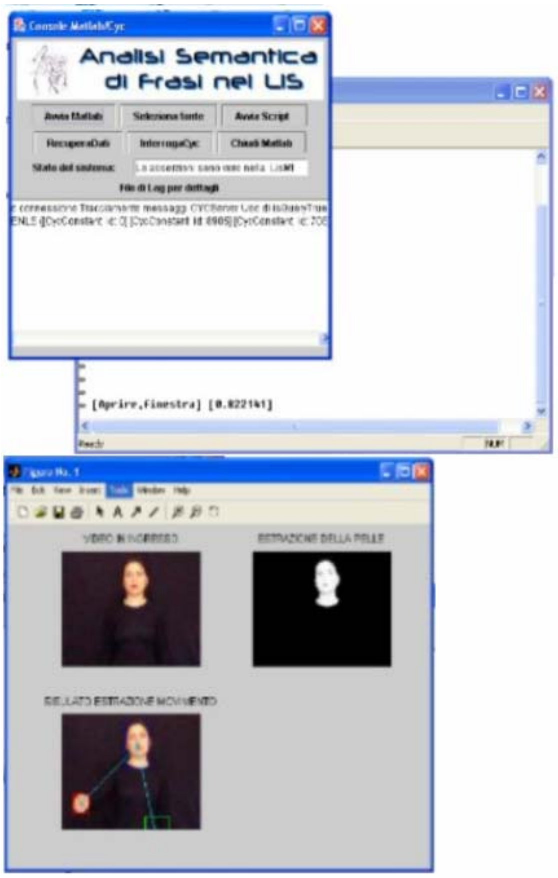


Figure 4: The Graphical User Interface of the system

The sign segmentation subsystem is based on the lips movement and uses a threshold on the width of the lips area of segment the sentence.
As mentioned before, we used OpenCyc capabilities to implement the reasoning module: we defined a microtheory called LisMt that is a specialization of AgentMt. The other modules are able to interact with the Cyc server by a suitable Java application that is executed in the Matlab workspace.
The defined a microteory LisMt includes 40 signs, and represents the current context. The first pre-processing step eliminates adjectives that usually are not relevant to investigate the correctness of the sentence. In the following an example of processed sentence is reported. The true signs are “ROBOT LIBRO BIANCO PRENDERE” (robot take white book).
In the example the adjective bianco (white) is not considered in the first evaluation: it will be considered if no possible correct sentence is founded, and in this case the system will search for a possible substitute as it will be shown in the next example. The syntactic analysis and the consistence with the local context are obtained generating the following query:

(#$and(#$relationAllExists #$performedBy #$prendere #$robot)
(#$relationAllExists #$objectActedOn #$prendere #$libro))

The Cyc server checks if
- #$robot is an actor
- #$prendere is a verb
- #$libro is a object

the sentence is compatible with LisMt and returns TRUE and the sentence is accepted. Now, we describe a sentence with an error: robot cucinare vestito [robot dress cooks]. The query generated is

(#$and (#$relationAllExists #$performedBy #$cucinare #$robot)
(#$relationAllExists #$objectActedOn #$cucinare #$vestito))

and Cyc returns FALSE, because there is not the context validation:
- #$robot is an actor
- #$cucinare is a verb
- #$vestito is a object
- but vestito is not a #$Food and cannot be cooked

From the analysis of the SOM maps we see that vestito was the first recognized sign, but pasta – the correct one – was the second with a similar rank index in the ranked list of possible meanings. The sign cucinare was recognized with a consistent difference from the second most probable. Then, the system tries to investigate the possible substitutes of vestito. A query with a free variable is generated:

```
CycList error_query= cycAccess.makeCycList(“ (#$relationAllExists #$objectActedOn #$cucinare ?X)”);
CycList substitutes_list= cycAccess.askWithVariable( error_query, new CycVariable(“?X”),mt);
```

The returned list is {pane, pasta}, i.e. all the objects of the current context that can be cooked. Pasta is selected because is included in the SOM candidates. The whole process, from video acquisition to sentence recognition, takes less than 1 second plus the movement duration (typically ~4 seconds for a complete sentence): the delay is mainly dues to the wait of the answer from OpenCyc.

Correct segmented sentences	76 (95.5%)
Correct translated sentences	66 (82.5%)
Erroneous translated sentences	10 (12.5%)

Table 1. Experimental results.

A test set of 80 videos of sentences has been used to check the system performances. We have obtained the 95% of correct segmented sentences (76 of 80). Moreover, 66 of this 76 sentences has been correctly translated, obtained a final success rate of 82,5%.

5. CONCLUSIONS

We have proposed a complete framework for sign language recognition that integrates a common sense engine in order to deal with sentences. The proposed architecture allows modelling and managing the knowledge of the recognition process in a simple and robust way. Moreover, the introduction of the semantic context resolves the problem of the analysis and validation of a sentence.

The presented experiments show that the system maintains the recognition rate high when the set of sign grows, correcting erroneous recognized single sign using the context. Table 1 shows the experimental results using 80 videos of sentences using 40 signs. Experiments demonstrate the goodness of the proposed approach. Future research will deal with the extension of the number of signs, allowing to use the system in more general contexts.

6. REFERENCES

Bauer, B., Hienz, H., Kraiss, K.F. (2000). Video-based continuous sign language recognition using statistical methods. In proceedings of the Intl. Conference on Pattern Recognition, vol. 2, pp. 463-- 366.

Bowden, R., Windridge, D., Kadir T., Zisserman A, Brady M. (2004). A linguistic Feature Vector for the Visual Interpretation of sign language. In Tomas Pajdla, Jiri Matas (Eds), proc. of 8th European Conf. on Computer Vision, ECCV04, LNCS3022, Springer Verlag, vol. 1, pp. 391--401.

Charanyapan, C., Marble, A. (1992). Image processing system for interpreting motion in the American Sign Language. Journal of Biomedical Engineering, vol. 14(5), pp. 419--425.

Cui Y., Weng, J. (2000). Apparence based hand sign recognition from intensity image sequences. Computer Vision and Image Understanding, May 2000, pp. 175--176.

Erdem, M. U., Sclaroff, S. (2002). Automatic detection of relevant head gesture in American Sign Language communication. In proceedings of the International Conference on Pattern Recognition, pp. 460--463.

Grobel, K., Assan, M. (1997). Isolated sign language recognition using hidden Markov models. In proceedings of the IEEE Internationak Conference on System, Man and Cibernetics, vol. 1 (12-15), pp. 162--167.

Hamilton, J., Micheli-Tzanakou, E. (1994). Alopex neural network for manual alphabet recognition. In proceedings of the IEEE Conference on Engineering in Medicine and Biology: Engineering Advances: New Opportunities for Biomedical Engineers, pp. 1109--1110.

Liang, R. H., Ouhyoung, M. (1998). A real-time continuous gesture recognition system for sign language. In proceedings of the International Conference on Automatic Face and Gesture Recognition, pp. 558--567.

Kim, J.S., Jang, W., Bien, Z.N. (1996). A dynamic gesture recognition system for the Korean sign language (KSL). IEEE Transaction on System, Man, and Cybernetics: Part B, April 1996, pp. 354--359.

Kohonen, T. (1997). Self-Organizing Maps, Springer-Verlag, NY.

Ma J., Gao, W., Wang, C., Wu, J. (2000). A continuous Chinese sign language recognition system. In proceeding of the International Conf. on Automatic Face and Gesture Recognition, pp. 428--433.

Miikkulainen, R. (1990). Script recognition with hierarchical feature maps. Connection Science, Vol. 2, pp. 83--101.

Starner, T., Pentland, A. (1995). Real time American Sign Language recognition from video using hidden Markov models. In proceedings of Symposium on Computer Vision, pp. 265--270.

Stokoe, W. (1978). Sign Language, Silver Spring, Linstok Press.

Sweeney, G.J., Downton C, (1997). Sign language recognition using a cheremic architecture. In proceedings of the Sixth International Conference on Image Processing and Its Applications, vol. 2(14-17), pp. 483--486.

Triesch, J., von der Malsburg, C. (1996). Robust classification of hand postures against complex backgrounds. In proceedings of the International Conference on Automatic Face and Gesture Recognition, pp. 170-175.

Yang, M.H., Ahuja, N., Tabb, M. (2002). Extraction of 2d motion trajectories and its application to hand gesture recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, Aug 2002, (pp. 1061--1074).

Vogler, C., Metaxas, D. (2001). A framework of recognizing the simultaneous aspects of American Sign Language. Computer Vision and Image Understanding, n. 81, (pp. 358--384).

Volterra, V. (1987). La lingua italiana dei segni: La comunicazione visivo-gestuale dei sordi. Il Mulino, Bologna.

Waldrom, M.B., Kim, S. (1995). Isolated ASL sign recognition system for deaf persons. IEEE Trans. on Rehabilitation Engineering, vol. 3(3), pp. 261--271.

Wang, C., Gao W., Shan, S. (2002). An approach based on phonemes to large vocabulary Chinese sign language recognition. In proceedings of the International Conference on Automatic Face and Gesture Recognition, (pp. 393--398).

Wilson, E. J., Anspach, G. (1993). Applying neural network developments to sign language translation. In proceedings of the IEEE-SP Workshop on Neural Networks for Processing, (pp. 301--310).

The methodological, linguistic and semiological bases for the elaboration of a written form of French Sign Language (LSF)

Brigitte Garcia

UMR 7023 *Structures Formelles du Langage*,

Université Paris 8, 2 rue de la Liberté, Saint-Denis (France) & CNRS

E-mail: bridge.garcia@wanadoo.fr

Abstract

This paper takes up the question of elaborating a graphical representation for French Sign language (LSF), beginning with the specificities of the socio-cultural context in which this question arises for those most directly concerned, that is the Deaf. We underline especially the vigilance required, when confronted with the influence of the written form of the Vocal language on linguistic (and therefore graphical) representations of Sign language (SL). We then present the results of a field survey, which allow us to justify and define our objective: write and not transcribe LSF. Next we explain precisely how the admitted limitations of current graphical systems for SL call into question the validity of the principles of segmentation adopted by consensus, which results from the influence of model of dominant alphabetical writing systems and of the focalisation only on lexical signs taken out of context, at the expense of the structural specificities of SL. We present on these bases the major principles of the alternative method begun for LSF, based on the descriptive model proposed by Cuxac (2000). We wish in particular to explore and evaluate his hypothesis of low-level morpho-phonetic segmentation, thus opening the way for an at least partial morphemographic representation.

The question of electronic representations of sign languages (SL) is of great importance not only for the recognition, the dissemination and linguistic and cross-linguistic study of these languages but also for the constitution of available resources to teach them. This digitalised representation brings into play all that pertains to the video capture of SL and of its treatment by computer. This raises however the closely related question of their representation in a graphical form, quite a different problem from the preceding one, since it implies the elaboration of a meta-language translated into another modality (here, the graphical modality). The present paper takes up this last point, and to do this relies on reflections and research carried out on these topics in the last few years in the framework of national projects and international cooperation ((Garcia & Boutet, 2003, 2006).

In order to determine the linguistic and semiological implications of the problem of elaborating a graphical formalisation for SL, and here especially, for LSF (see sections 3 and 4), two preliminary tasks are indispensable. First of all, we must recall the specificities of the historical, social, cultural and educational context in which the question, for the Deaf, of providing their languages with a graphical form is situated (section 1). These specificities in fact allow us to take stock of the preliminary conditions and demands, which are as much linguistic and methodological as they are deontological, and which are required in order to perfect the evaluation of graphical representations for these languages. The second task that is necessary to help in determining the semiological and formal choices of the projected graphical system consists in specifying which functions we wish to assign to such graphical representations (section 2).

1. The Specificities of the context in which the question of elaborating a graphical representation for SL is raised

No writing system of any language has ever been elaborated by anyone other than the speakers of the languages in question. In this sense, it is difficult to imagine, that a written form of SL could be elaborated without a close collaboration with those who use these languages. Several characteristics of the linguistic and political situation of the Deaf render the addressing of this evident truth, however, rather complex.

We know the consequences for a language due to the fact of its not disposing of a written form, in terms of status, of political and cultural recognition and in terms of its power of dissemination. For the institutional SL practiced by the Deaf in Western-World societies, these consequences are still greater. As they are languages without geographical specific definition, and their political recognition remains quite fragile, these SL are obliged to exist side by side with languages that are in reality doubly dominant: they are vocal languages (VL) and they dispose of a written form which is the unique national written language, so that this written form constitutes for the Deaf the only means of gaining access to information and to knowledge. This problem is compounded in the vast majority of these countries by an additional difficulty for the Deaf, who are in most cases deprived of any reasonable access to the written form of the national language. For France, the only figure available estimates the level of illiteracy at 80% among the Deaf (Gillot, 1998). These enormous difficulties, observed in most countries having an institutional SL, are above all the result of a choice made in most cases decades before by educational authorities to make use not of SL – the only natural language – but of the national VL as instructional language for the education of the Deaf (e.g. Chamberlain *et al*, 2001).

This reality has two crucial consequences for our undertaking. First of all, it has produced a very ambivalent relationship on the part of the Deaf toward the written form of the VL, and moreover, to the written word in general. A survey that we recently carried out in France with Deaf users of LSF (Garcia & Boutet, 2006) has shown among other things that the Deaf feel a form of oppression, of forced dependence and have a mental block toward written French and toward written forms in general; but they simultaneously have a tendency to consider the particular alphabetical and linear form that it takes for the VL as sacred, and see in it the only possible form of writing. Another serious consequence is that the absence of any mastery of the written form on the part of Deaf signers makes it difficult for them to gain access to linguistic knowledge and accumulated reflections on languages in general, and on SL in particular, as for instance the history and the semiology of writing systems. This alone would yet enable them to step back from their own language and from the written forms, which is necessary for the elaboration of an adequate graphical meta-language.

Conversely, we must stress the consequences that its written form has on any language, which affect our manner of perceiving it and of describing it and which consequently could affect its evolution. As regards VL, many studies have shown to what extent the standards for spoken forms are permeated by those of written forms, and how on a meta-linguistic level, the written form conditions the perception that linguists have of spoken forms and, as a result, the description that they give of them (e.g. Blanche-Benvéniste, 2000). One could go even further by insisting, as did Stokoe (1991), on the fact that the conceptual foundations of general linguistics have been to a large extent elaborated from the study of the written forms of VL. The elaboration of a graphical form for SL, which are profoundly and literally *face to face* languages, is therefore for these languages and for those who use them, not a trivial undertaking. On the contrary it can very strongly orientate the representations, the description and the very evolution of these languages. The particular political and social condition of SL implies an additional risk: the influence the dominant written form of the national VL can exert on these socially fragile languages. In fact SL are not exempt from all forms of graphical representation, and the noteworthy point here is precisely the influence, hitherto recognised, of the written forms of VL on these languages and on their description, through the choices of graphical representations made to date for the SL. I will return to this point later, limiting myself here to stress the particular implications of such influences for these languages operating in another modality, and which in fact display very strong structural specificities.

It therefore seems to us that two requirements should be stated as the basis for any undertaking that purports to elaborate a graphical form for SL. On the one hand, there is the setting in place of collaborative structures to create the conditions for a real and priority involvement of Deaf signers in the process of reflection on the graphical representation of their language[1]. On the other, particular care must be taken as to the exact correspondence between these graphical forms and the structural specificities of SL, and as regards the levelling down of these specificities that can be induced, directly or indirectly, by the social et cultural dominance of the written forms of VL.

On this basis, the other major question is that of the needs which should be addressed by a graphical formalisation of SL. What functions should it fulfil? It is obviously essential to answer this question, in order to determine the linguistic and semiological choices to be made, that is, the form of the graphical system itself.

2. The functionalities of a graphical system for SL

Any notation indeed, regardless of the form it takes, is evidently not a simple reflection of language in absolute terms. It is based on the integration of a certain number of more or less explicit theoretical hypotheses about what structures this language, but also on the formal choices (choices notably in the level of analysis and of encoding), which both depend on the potential uses and users of the system. So these uses and users have to be identified.

2.1 Transcribe

The first essential type of function that a graphical system of SL should assume is inherent in the linguistic description and is precisely what has dominated nearly all notation systems elaborated up until now for SL. These are functions that answer to the needs of researchers, as much for the constitution of dictionaries as for the transcription, the preservation and the exchange of corpus data for the SL under study. The issues at stake here are of a specific nature. The transcription should indeed be representative of the phenomena observed by the linguist, and it must make the structures that he brings to light (and which he has hypothesized) “appear”, at the particular level of analysis that he finds of interest (phonetic, phonological, morphological, syntactic, or discursive) and in accordance with his point of view and objective. Due to its heuristic aim, the transcription system should also allow the researcher to note, at the level of analysis he has chosen, as many details as possible – even those details that he perhaps was not expecting. In the perspective of exchanges between researchers and especially for cross-linguistic research, it is equally important that these graphical representations allow for the exact mental reconstitution of the language data transmitted in this way.

This very specific function of transcription is however not the only one that a graphical system for SL must fulfil. One of our first questions was in fact to know whether there were, in the community of Deaf signers of LSF (who had never been consulted on this point), needs relative to the specific graphical representation of their language. This was the reason for our qualitative survey mentioned above. For us it was a matter of evaluating -together with Deaf signers - whether and by what means the cognitive, social and cultural functions of writing are fulfilled for these Deaf citizens of societies based on the written word, and, if

[1] On this account, the *LS Script* project includes as one of its partners the IRIS association, which brings together Deaf teachers working in bilingual educative structures (Toulouse).

they are not, whether and how they could be.

2.2 But also and above all to write

This investigation forms the foundation for the totality of our work toward a written form of LSF. It has made it possible to directly involve the French Deaf community in the process of reflection, and also to bring up the fact that there are a number of situations in which, in fact, neither written French—even when it is mastered— nor video are regarded as being satisfactory. Our first major observation is this: Deaf signers (including those who could be considered as illiterate) dispose of very many specific graphical practices, the aim of which is to notate LSF— regardless of how well they master written French. Here it is exclusively a matter of practices for oneself, or between members of the Deaf community. The one encountered most frequently, even if one masters written French, is a specific use of the latter, that those interviewed designated in terms of either "*LSF-French*" or "*written LSF*". This involves the lining up of French words, following —according to them — the syntax of LSF and what they refer to as "sign-words". The weaker their mastery of French, the more this LSF-French is mixed with drawings, until it becomes literally based on drawings and other graphical symbols. This can lead to the development of completely original graphical systems, often individualised and more or less standardized, from which any "sign-word" is absent. We will insist in what follows on those situations in which the Deaf who we encountered have recourse to such specific practices even when they master written French, since, they claim, the latter is ill-adapted to their needs. These practices correspond to as many functions as could be assigned to a specific graphical form.

First of all, this concerns situations in which one has recourse to graphics in its primary function as support for the construction of thought, for example to prepare a production in SL (a conference or an appointment). Most often, and especially for those who have only some, little or no command of French at all, one has recourse to the continuum mentioned above, which runs from «*LSF-French*» to the exclusive use of idiosyncratic graphical symbolisations. In any case, as the thought process functions in LSF, this situation is presented as being one of the most frustrating —while even this use of the written word as a support for a cognitive elaboration corresponds to one of its key functions. Another problematic situation occurs with any written support for an "oral" presentation in LSF, the equivalent of notes for a hearing conference speaker. The particular difficulty is then that the utilisation of written French as a support disturbs the fluency of the production in LSF and incites to produce signed French. Both problems that we have just mentioned are to be found, according to the teachers we encountered, in the instruction of LSF in a school setting, where the pupils, for exercises of self-correction or for evaluation, have to prepare and then give a production in LSF in front of the camera.

We will mention two other problematic situations: first, that of note-taking from a course, a conference or a meeting in LSF, any recourse to French often being deemed inadequate, and even more so if one wants to conserve a specific formulation in LSF. Then there is the case, moreover, where one creates a literary or artistic type of production in LSF, of which one would like to retain the graphical expression, allowing one to memorise it and conserve it, but also to re-work it in detail — which the video format is not flexible enough to allow for. It is a case such as this which has given rise to the most systematised graphical inventions, even if most often, their use is limited to a very few persons.

According to those interviewed, at least a part of the functions of the written word are for SL, already fulfilled or about to be by video and the new visual technologies—which they often consider as the "written form" of LSF. However, to a majority of them, there are currently certain limitations for video and the new visual technologies that will never be extended. First, as regards the access to SL data banks (search engines): the predicted potential in the medium term concerning digital image recognition (for movement and form) are still far from being equal to the economy of means inherent in the formulation of query *via* a specific graphical representation of language data being sought. More fundamentally however, for the majority of persons encountered, video presents intrinsic limitations which prevent it from specifically playing the role of support for the elaboration of a reflection, and for which as we have seen, written French, even when mastered, is ill-adapted. This is due firstly to the maintaining of the visible physical presence of the signer: video, by this very fact, remains bound to the face to face communication; it forbids, above all, the distancing that the written form authorises, *a fortiori* when it's a matter of one's own image. It is secondly due to the fact that video, through its streaming in time, does not allow for the simultaneous view of what is being recorded and what has already been recorded — a simultaneity which yet serves as the basis for the inherent potentials of writing.

These various observations argue for the elaboration of a graphical system that allows for written production in LSF, and for this reason, the development of a writing system is the prior objective of our *LS Script* project. From this point we must insist on the fact that here it is a question of quite a different function that that of transcription mentioned above. As a mode of communication on its own and having specific functions itself, a writing system allows for the direct production of propositional content, and in reception, a direct access to meaning which does not issue from an oral production beforehand. The function of a transcription system is on the other hand to graphically represent utterances first produced in an oral form, spoken or signed. The difficulties encountered by linguists who describe spoken forms of VL to transcribe their corpora are a true reflection of the gap that exists between writing and transcription. Indeed, writing systems only manage to fulfil the function of transcription imperfectly, and at the price of a great compromise of conventions (Blanche-Benvéniste, 1997; Habert *et al*, 1987).

Keeping in mind this necessary distinction between writing and transcription which has rarely been taken into consideration, it is helpful to analyse the linguistic and semiological choices that were made while devising the systems of notation and of annotation that have already been elaborated for SL, and then to address the problems such an analysis can bring to light.

3. Questions raised by existing systems of notation and of annotation for SL

A good many graphical systems have been specifically created for SL. In the great majority however they have been conceived by researchers in response to their own research needs[2]. These are transcription systems that can be divided into two main types: on the one hand, there are autonomous systems, *i.e* systems of notation based on rules and on particular modes of representation requiring no knowledge of another written form (e.g, Bébian, 1825; Stokoe, 1960; HamNoSys, 1989)[3]; on the other hand, there are systems of annotation, characterised by the fact that the medium of representation takes a pre-existing written form, that of the national VL (*e.g*. Johnston, 1991; Bouvet, 1996; Cuxac 1996).

3.1 The Limitations of current systems of notation and of annotation

Several recent studies have stressed the limitations of these two groups of systems for transcription itself as well as for writing (Pizzuto & Pietrandrea 2001□Garcia & Boutet 2003, 2006). The greater part of notation systems proceed, more or less explicitly, from the analysis of conventional signs taken out of context —and notably, out of their spatial context— into parametric elements from their visual aspect and rendered in linear form, the selected parameters being more or less those identified by Stokoe (1960 and 1965) following, moreover, the analysis *princeps* of Bébian (1825). It is hardly surprising then that they permit a readable and representative graphical restitution neither of simultaneous phenomena, nor of the internal variations of signs in discourse nor, more generally, of the phenomena of the spatializing of semantic and grammatical relations. Annotation systems, devised to make up for these limitations, only manage in reality to *describe* these discursive phenomena through the recourse to the written form of the VL. For the graphical expression of lexical units, these systems can integrate existing notations and/or avoid the problem of any notation whatsoever of the internal structure of these signs and of their variations in discourse through recourse to the conventional principle of glossing[4]. What the analysis of present-day systems of notation and annotation of SL finally reveals is the existence of a hiatus between on the one hand, notations centring on the unit of the sign and inadequate for the recording of discourse, and on the other, systems of annotation for discourse in which the recording of these signs in context is relayed (masked) by the recourse to the written form of the VL (gloss). This hiatus, in itself, calls into question the relevance of the segmenting principles that have been adopted.

3.2 Indirect forms of influence by the written forms of VL

It seems to us that two essential problems are able to account for these limitations. The first concerns the influence exerted indirectly by the written forms of VL. If we remember that the historical diversity of writing systems of VL themselves results from the typological diversity of these languages, it is difficult to see how the models for the written form of VL could be suited, as such, to languages as typologically different as SL – unless we consider that languages of the audio-oral modality have exhausted the totality of all possible graphical forms. However, the fundamental semiological principles of existing systems of notation for SL, *i.e* the encoding of the mere formal aspect of units considered as the equivalents to phonemes and linearization, result from a very direct adaptation of alphabetical principles elaborated for VL. They thus completely overlook what is one of the most marked structural specificities of SL, that is, the spatializing of nearly all semantic and grammatical relations. The one notation system that has truly innovated in terms of the semiological exploitation of the graphical modality is the *Sign Writing* system (Sutton, 1999), *via* its direct utilisation of graphical surface as an *analogon* of the space in front of the signer and the exploitation of the position of symbols on this surface as referring analogically to the relative position of bodily articulators thus symbolised. One of the aspects of our work consists in a systematic exploration of the semiological potential inherent in the visual-graphical modality, notably in the two areas of the notation of non-linguistic bodily movement and of phenomena exploiting topological spaces (conventions utilised in molecular chemistry).

The other problem is concentration of inventors of notation systems on lexical signs. The difficulty here owes not only to the fact that they most often begin from the analysis of these signs outside of discursive context, but equally to the fact of considering them, on the one hand, as the principal if not the only linguistic unit of SL, on the other, as the ultimate unit of meaning. We can see in this exclusive concentration an indirect form of influence of writing alphabetical systems (the “word”). The theoretical model we have adopted, that of Cuxac (2000, 2004) allows us to pose the question differently. This model notably integrates at the very centre of its preoccupations number of (signed) productions which are, in part, usually considered in the literature as not belonging to SL (*i.e* as “ungrammatical”, e.g Liddell, 2003) – whereas they are extremely frequent, especially in story-telling (see Sallandre, 2003). These structures, which Cuxac calls “highly iconic structures” (HIS), are characterised by a very strong iconicity and by the fact they include only little or no lexical signs (“standard□ signs”).

On these grounds and unlike earlier approaches, we choose to *start from* the structural specificities of SL and

[2] *Sign Writing* (Sutton, 1999) constitutes in this sense a notable exception.

[3] See (Miller, 2001) for an inventory

[4] On the various problems posed by this recourse to glosses, *i.e* to words of the written VL and, especially, the way in which they distort the representations and the description of SL, see (Cuxac, 2000), (Pizzuto & Pietrandrea, 2001) and (Pizzuto *et al* in the present proceedings). It clearly appears that working towards a written form of SL is presently the best means of making the transcription systems themselves progress.

to reckon on the necessity, as well as the possibility of imagining alternative graphical ways. This by no means excludes the recourse at some stage to the range of semiological graphical processes put in place in the written forms of VL —other than phonographematic ones.

4. The initial lines of inquiry: re-thinking low-level segmentation

To start from the specificities of SL —at least LSF— considered in the framework of Cuxac's model, which seems to us to reproduce them most exactly, means that this formalisation should concentrate on two key aspects. On the one hand, an alternative investigation of low-level structuring, and on the other hand the modelling of phenomena appertaining to a pertinent utilisation of space. In this paper, I will only develop the first point, which has been the object of our most extensive investigations to date. As regards the graphical formalisation of spatial phenomena, and particularly, the constructing of reference and processes for the constructing of co-reference (anaphora), I will limit myself here to a remark. The graphical system to be elaborated does not aim at a representation of previous oral productions that it would simply transfer to a graphical form, like some visual anamorphosis in two dimensions of three- or four-dimensional phenomena. The true semiological challenge is to succeed in elaborating a graphical *interpretation* that respects what is structurally relevant in SL. In this sense, the original semiological choice made by *Sign Writing* to simply represent a visual phenomenon by another visual form, that is to exploit 2D graphical space as a "flattened" *analogon* (lacking depth) of the signing space, seems to me to be also one of its limitations, especially for the restitution of sophisticated phenomena concerning the spatializing of *loci*, and of anaphora.

The issue at stake in low-level modelisation is to determine the principles legitimating the choice of graphical units. In their great majority, current descriptions of SL propose a phonological type of modelisation, whether the reference would be that of the new phonological theories or that of a functional type of phonology. It is to this latter model and to the "phonetics" that it implies that existing notation systems including *Sign Writing* refer— whether explicitly or not. Here the lexical signs are analysed according to purely formal parameters aiming to explain their visual form (configuration, orientation, location, movement, ± facial expression), these parametric units being assimilated either to phonemes or to phonetic units. Beside the admitted awkwardness of these notation systems, the theoretical motivations for calling these parametrical principles into question are many. One of them, long remarked (Studdert-Kennedy & Lane, 1981; Jouison, 1995), is the difficulty encountered when trying to assimilate these parametrical elements to phonemes, since many of them carry meaning. Another difficulty is precisely due to the limited framework of the original source of these parametrical elements, which is the "standard sign". In the perspective of a description of LSF that places the HIS at the heart of the model, another important type of minimal unit has to be taken into account: units of "transfer"[5], a part of whose constitutive elements are non discrete.

The hypothesis defended by Cuxac is based on a low-level structuring that is not phonemic, but from the outset, morphemic, with minimal elements resembling bound morphemes; it opens up an important alternative route to the modellisation of SL as well as for their graphical formalisation. This hypothesis calls for the morphemic compositionality of standard signs, analysed as minimal units of production comparable to molecules compounding atoms of meaning, elements that are partially commutable but non autonomous. This hypothesis remains to be validated. It is toward this direction that our work is orientated, which consists in taking the inventory, parameter by parameter —over the totality of the LSF lexicon presently accounted for (Girod *et al*, 1997)— of the lowest level morphemic components, and by ranking their values. The objective is to establish the productivity of these morphemic elements and to identify any possible rules of compositionality. The analysis, which is under way, has been concerned up until now with the configuration, the location and the direction of movement. It confirms the existence of a strict organisation of morphemic values and the coincidence of the most productive among them with those attested as constituents of HIS. It allows us moreover to highlight a number of cases of interdependency between parameters, and in particular, between the configuration and the direction of movement (Boutet, 2005; Garcia & Boutet, 2006).

For us, such a modelisation is doubly interesting. In part at least, these morphemic elements seem common to standard signs and to HIS: identifying what then would be form-meaning constants would put (us) on the trail of graphemes transversally common to two types of structures, HIS and standard signs. Moreover, a morphemic analysis opens up the possibility of an at least partially morphemo-graphical notation and no longer merely formal, and motivates the recourse to certain combinatory methods exploited by this kind of writing systems: the association of "phonetic" determinatives (ideo-phonograms) or semantic determinatives to these morphemograms (combinations of morphemograms in the manner of Chinese syllogigrams). The demonstration of systematic interdependency between parameters should permit besides to limit the final number of units to be retained for the notation.

One of the challenges of this long and exacting work (of which I only give a glimpse here) directly concerns one of the topics covered in this conference, that I will mention to finish with; that is the constitution of dictionaries for SL. The form taken by these lexicons and dictionaries that have been elaborated since the 18th century is a direct result of the absence of a historical written form for these languages[6]. For a number of them,

[5] Cuxac discerns three major types of transfers, which constitute the HIS: personal transfers (the signer "becomes" one of the actants of his utterance), transfers of form (which allow one to describe any type of form) and transfers of situation (which allow one to represent the movement of an actant in relation to a stable localising referent).

[6] As regards LSF, we refer the reader, for an analysis and a very exhaustive inventory of these dictionaries, to the work of

and at any rate for the only dictionary presently existing for modern LSF (Girod *et al*, 1997), the method of classification and of description consists in fact in associating a given sign in LSF (possibly represented by a drawing) with one or several words in written French. Jouison (1995) pointed out the simplistic character of such a representation, which he considered as an indirect form of influence exerted by the written form of the VL, and which focussed the attention (especially of linguists) on this one level of analysis. Cuxac (2004) stresses the problem posed by these dictionaries: where the dictionaries of VL propose an average of 50,000 entry words for these languages, the dictionaries of SL offer at best 5,000. Rather than come to a conclusion as to the lexical indigence of SL, one must once again question the structural relevance of the method of classification, which does not take into account the true bases for the structural organisation of SL. The alternative solution would be a system of entry words by morphemic element[7], and if at all possible by morphemo-graphical element.

References

Bébian, A. (1825). *Mimographie, ou essai d'écriture mimique, propre à régulariser le langage des sourds-muets*. Paris, L. Colas, in-8°.

Blanche-Benvéniste C. (2000), *Approches de la langue parlée en français*. Ophrys, coll. *L'essentiel français*, Paris.

Bonnal, F. (2005). Sémiogenèse de la langue des signes française : étude critique des signes attestés sur support papier depuis le XVIII[e] siècle et nouvelles perspectives de dictionnaires. Thèse de doctorat, Université Toulouse Le Mirail.

Boutet, D. (2005). Pour une iconicité corporelle. In Actes du colloque TALN RECITAL de Dourdan, pp. 345-348.

Bouvet, D. (1996). *Approche polyphonique d'un récit produit en langue des signes française*. Presses Universitaires de Lyon.

Cuxac, C. (1996). Fonctions et structures de l'iconicité dans les langues des signes. Thèse de Doctorat d'État, Université Paris 5.

Cuxac, C. (2000), *La Langue des Signes Française (LSF). Les voies de l'iconicité*, Ophrys, Faits de Langues, Paris.

Cuxac, C. (2004). 'Phonétique' de la LSF : une formulation problématique. In Linguistique de la LSF : recherches actuelles , Silexicales 4, A.M. Berthonneau et G. Dal (dir.), Université de Lille 3, pp. 93-113.

Chamberlain, C., Morford, J.P. & Mayberry, R.I. (eds) (2000). *Language acquisition by eye*, Lawrence Erlbaum Associates publishers, Mahwah, New Jersey, Londres.

Garcia, B. & Boutet, D. (2003). Vers une formalisation graphique de la Langue des Signes Française (LSF). In *Langue des Signes Française (LSF)□ enjeux culturels et pédagogiques*, C. Cuxac ed. *La nouvelle revue de l'AIS*, CNEFEI, 23, pp. 49-62.

Garcia, B. & Boutet, D. (2006). Finalités et enjeux linguistiques d'une formalisation graphique de la langue des signes française (LSF). In Les langues des signes (LS): recherches sociolinguistiques et linguistiques, n°7, DYALANG, http://www.univ-rouen.fr/dyalang/glottopol/numero_7.html.

Gillot, D. (1998). Le droit des sourds, governmental report to Lionel Jospin.

Girod, M. *et al.* (1997). *La langue des signes, dictionnaire bilingue LSF/français*, t. 2 et 3. IVT Éditions, Paris.

Habert B., Nazarenjo A., Salem, A. (1997). *Les linguistiques de corpus*, Paris, Armand Colin.

Johnston, T. (1991). Transcription and glossing of sign language texts: examples from Auslan (Australian Sign Language). In *International Journal of Sign Linguistics*, Vol. 2, n° 1.

Jouison, P. (1995). *Écrits sur la Langue des signes française (LSF)*, B. Garcia (ed). Paris, L'Harmattan, coll. « *Sémantiques* », Paris.

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

Miller, C. (2001). Some reflections on the need for a common sign notation. In Wilbur R. (éd.), Sign Language & Linguistics, vol. 4 n° 1/2, John Benjamins Publishing Company, Amsterdam / Philadelphia, pp.11-28.

Pizzuto, E. & Pietrandrea P. (2001). The notation of signed texts. Open questions and indications for further research. In Wilbur R. (ed), *Sign Language & Linguistics*, John Benjamins Publishing Company, Amsterdam / Philadelphia, vol. 4 n° 1/2, pp. 29-45.

Prillwitz, S., Leven, R., Zienert, H., Hanke, T., Henning, J. *et al.* (1989). *Hamburg Notation System for Sign Languages. An introductory guide,* HamNoSys version 2.0. Hamburg: Signum Press.

Sallandre, M. A. (2003). Les unités du discours en Langue des Signes Française. Tentative de categorisation dans le cadre d'une grammaire de l'iconicité. Thèse de doctorat, Université Paris 8, Paris.

Stokoe, W. C., (1991). Semantic Phonology. In *Sign Language Studies,* W. C. Stokoe (ed). Lindstock Press.

Stokoe, W. C., Casterline, D., Cronenberg, C. (1965). *A Dictionary of American Sign Language, based on linguistic principles*. Gallaudet College Press. Washington D.C.

Stokoe, W. C. (1960), Sign Language Structure, An Outline of the Visual Communication Systems of the American Deaf ». In *Studies in Linguistics*, occasional paper n°8, Department of Anthropology and Linguistics, University of Buffalo.

Studdert-Kennedy, M. & Lane, H. (1981). Clues from the differences between signed and spoken language. In *Signed and spoken language: biological constraints on linguistic form*. Ursula Bellugi & M. Studdert-Kennedy (eds). Harvard University Press.

Sutton, V., (1999). *Lessons in Sign Writing. Textbook & workbook* (2nd edition, 1st edition 1995), La Jolla, CA: Deaf Action Committee for Sign Writing.

F. Bonnal (Bonnal, 2006).

[7] Bonnal (2006) proposes, on the basis of Cuxac's hypothesis, a sample of what such a dictionary for LSF could be like.

Glossary compilation of LSP including a signed language: a corpus-based approach

Cynthia Kellett Bidoli

Scuola Superiore di Lingue Moderne per Interpreti e Traduttori – Università degli Studi di Trieste
Via Filzi 14, 34127 Trieste, Italy
jkellett@units.it

Abstract

Sign language interpreters not only work in a 'community' context but also are called to conferences on deafness-related issues containing language for special purposes (LSP). In Trieste, within an Italian national research project, one particular area of research has been centred on investigating textual recasting that may take place during English to Italian Sign Language (LIS) interpretation based on the compilation of parallel multimodal corpora in English, Italian and LIS. Electronic analysis of the corpora enabled the collecting and concordancing of specialized terminology and the development of a pilot version of a trilingual electronic terminological dictionary (on CD-ROM). The glossary will contain dynamic imagery of LIS and will provide a useful and innovative tool for future interpreter trainees.

1. Introduction

Since 2002, joint research has been conducted by teachers of English at the universities of Turin and Trieste within two Italian national research projects on how and to what extent the English language influences cultural and linguistic communication in contact with Italian[1]. Italian society, as all societies throughout the world, conceals an invisible 'community of practice' within its confines composed of Deaf people[2], which (amongst other social categories identified by the research unit as belonging to different discursive domains[3]) was targeted by the Turin unit as representing a very intriguing area of interlinguistic/cultural contact to study wherein socially and ideologically marked identity traits are not only discursive characteristics but are negotiated in the very choice of communicative code: sign language - in this case Italian Sign Language (*Lingua dei Segni Italiana* - LIS). The access by Deaf cultures to international communication, has long been ignored in research in Italy in the field of English language/linguistics, and is considered in this particular research project as a relevant intercultural situation to investigate.

Research was first focussed on *Intercultural Practices and Strategies of Textual Recasting* to verify if and to what extent the production/reception of written and oral English discourse within a number of different domains leads to a propensity for cultural and linguistic intrusion from English into Italian and hence also into LIS (cf. Kellett Bidoli 2004, 2005b, forthcoming a; Ochse 2004a, b, 2005). Research has now turned to *Construction of Identity in Socio-political Discourse* to investigate the discoursal processes of construction, manifestation and negotiation of social identity, namely in intercultural situations such as the teaching of English to the Deaf or during interpretation from English to LIS.

In order to explain the link between the research projects, corpus linguistics and trilingual dictionary compilation, there follows a brief outline on contact between the English-speaking world and the Italian Deaf community.

2. The English language within the Italian Deaf community

Information about the English-speaking world reaches the Italian Deaf almost exclusively through written Italian sources: newspapers, magazines, translated books and articles, subtitled films and Italian websites on British or American issues. A minor visual source is provided by TV world news through simultaneously interpreted sign language at set times during the day (Kellett Bidoli 2004: 129). But there is also a good deal of direct exposure to the English language at school or university (Ochse 2001, 2004a), through the Internet, in the workplace, during periods of study abroad on cultural exchanges (Socrates Erasmus or Fulbright Scholarships), at public conferences on Deaf issues in the presence of English native speakers, and to a lesser extent at home (Kellett Bidoli forthcoming b). Deaf people are able to read and write in English if given adequate instruction at school, but where contact involves spoken English, professional interpreters are required to enable communication to take place. In 2003, a survey was conducted among professional Italian sign-language interpreters to determine the extent of English to LIS interpretation and discover which genres are commonly involved (Kellett Bidoli 2005a). It was found that interpreters with an active knowledge of English are more numerous than expected, but unfortunately only a handful are willing or able to mediate directly from English to LIS, and thus interpretation from English is normally always filtered in relay through Italian to a second interpreter who transfers the received message into a gestural/visual mode (LIS) for the deaf audience. The

[1] MIUR COFIN national project no. 2002104353 *Intercultural Discourse in Domain-specific English* and PRIN national project no. 2005109911 *Identity and Culture in English Domain-specific English* both coordinated by Professor M. Gotti. See: http:/www.unibg.it/cerlis/progetti.htm

[2] It is an accepted convention in the literature to use "deaf" (with a lowercase "d") to refer to the audiological condition, while "Deaf" is used to refer to those deaf people who share a sign language and distinct cultural values.

[3] The other discursive domains identified by the Turin unit are: diplomacy and on-line information/persuasion on socio/economic and human rights issues.

survey uncovered a number of genres within the context of conference interpreting and in particular the field of linguistics (conferences on various linguistic aspects of sign language and interpretation). This finding led to closer investigation of interpreted discourses and more specifically to the selection of an LSP corpus of representative, authentic, English discourses aimed at discovering to what extent the English language influences interpreted LIS (either directly or filtered through Italian) and hence the message received by a deaf end-user, and to what extent divergences may arise due to basic cultural or linguistic distinctions (Kellett Bidoli 2004, 2005b, forthcoming a).

3. From corpus to glossary

Video-recorded speeches were selected[4] to create a small corpus of 12,616 English tokens of which there were 3,075 types. The original video recordings in VHS were transformed into a digital corpus for electronic analysis and viewing of the interpreted signed language using **C**ode-**A**-**T**ext **I**ntegrated **S**ystem for the **A**nalysis of **I**nterviews and **D**ialogues software (C-I-SAID: Scolari, Sage Publications) capable of handling multimodal source data in the form of media files and plain text (sound, video and written text). The original sound files of the speeches in English were transcribed, together with the visual signed discourses to provide aligned, multimodal, parallel corpora to work on in order to reveal intercultural and linguistic aspects of textual recasting. The parallel corpora were composed of:

- a written transcription in English of the original spoken discourses;
- Italian glosses of the LIS signs (6,643 tokens and 1,819 types) transcribed with the help of an interpreter;
- a written 'interpreted' version in Italian of the signed corpus checked by a deaf teacher of LIS;
- a written 'interpreted' version in English of the signed corpus.

The horizontally aligned discourses in English and LIS were compared providing clear evidence of occasional disparity (from lexical items to whole chunks) leading to several instances of intercultural or interlinguistic communicative failure through semantic misrepresentation, distortion or omission. An example is given below which describes a preliminary exercise used in simultaneous interpretation training and illustrates corpora alignment:

ENGLISH: Students will listen to a fairy-tale that they know and they will…, they're asked to render this fairy-tale in their own words. They usually know the ideas of Little Red Riding Hood, let's say, anything that they know about. So they do not have to focus on the words, but they automatically grasp the concepts.

LIS Glosses: STUDENTE ASCOLTA, FAVOLA CONOSCERE, CONOSCERE POI IO DOMANDARE STUDENTI DOVERE PRODURRE PAROLE NO, MA DOVERE RACCONTARE, RACCONTARE QUALSIASI SAPERE FAVOLA, MA CONCENTRARE SOLO 1 PAROLA, PAROLA, PAROLA NO (+ *negazione con la testa*), DOVERE CAPIRE CONCETTO

(**STUDENT LISTEN FAIRY-TALE KNOW KNOW, THEN I ASK STUDENTS MUST PRODUCE WORDS NO, BUT MUST NARRATE NARRATE ANY KNOW FAIRY-TALE BUT CONCENTRATE ONLY 1 WORD WORD WORD NO** (+ *head negation*), **MUST UNDERSTAND CONCEPT)**.

Italian interpretation of the LIS: Gli studenti ascoltano una favola che conoscono bene. Dopo non chiedo loro di ripetere le parole, ma di raccontare quello che conoscono della favola. Non devono solo fermarsi sulle singole parole, devono capire i concetti.

English interpretation of the LIS: ***Students listen to a fairy-tale they know well. Afterwards I don't ask the students to repeat the words but to narrate anything they know about it. They mustn't focus only on single words; they have to understand the concepts.***

The most complex phase of the research involved the transformation of the sign language into a written form for electronic analysis. C-I-SAID was not designed specifically for sign language transcription and analysis. Since this research began, software for this purpose has been developed in the United States which can be applied to various languages[5]. Also not having access to a real-time sign language recognition system[6], the LIS was laboriously, manually transcribed into glosses with the assistance of both a professional LIS interpreter and a deaf teacher of LIS. The glosses were typed out according to sign language transcription conventions using capital letters and hyphenation, and initially appeared as a continuous string of words with no breaks or punctuation whatsoever. During a second phase the text was broken down into meaningful punctuated segments according to natural 'intonation' markers, pauses, and non manual communication: the gestures and facial expressions of the interpreter

A major problem was finding a solution to the transcription of the meaning conveyed by the non manuals: mouthed (not signed) words; nodding; raised eyebrows (as a question marker); rotation of the shoulders to the left or right (to identify agency); and so on. The

[4] W.C. Stokoe, a paper on the evolution of sign language "Hands, Eyes and Language", presented at the First National Conference on Sign Language, *Studi, esperienze e ricerche sulla lingua dei segni in Italia*, ENS, Trieste 13-15 September 1995, published in Italian in Caselli & Corazza 1997; W.P. Isham, "Research on Interpreting with Signed Languages", C.J. Patrie, "Sequencing Instructional Materials in Interpreter Education", and B. Moser-Mercer, "The Acquisition of Interpreting Skills", all three papers presented at the International Conference "Meeting of Sign and Voice", University of Trieste, Trieste 12-13 December 1997, published in Gran & Kellett Bidoli 2000.

[5] Such a system is SignStream, see: www.bu.edu/asllrp/signstream/contact.html.

[6] Hidden Markov models first used for speech and handwriting recognition have been adapted to the complex recognition of hand gestures. View-based gesture recognition can now recognize sentence level continuous American Sign Language using a single camera to track an ASL signer's hands. See: Starner, T., and Pentland, A., (1995) *Real-time American Sign Language Recognition from Video Using Hidden Markov Models*, International Symposium on Computer Vision.

glossed LIS version was thus filled with numerous bracketed annotations (rather than conventional straight lines with symbols placed above the glosses), as C-I-SAID automatically ignores all text in brackets, enabling rapid word counts and concordances to be performed.

On completion of the lengthy transcription of the LIS glosses, the text was segmented into small meaningful units, using the customary 'musical score' format for the transcription of bilingual mediation (as illustrated above). The addition of punctuation was also necessary in order to respect the requirements of C-I-SAID which is programmed to parse 'segments' of dialogue chronologically and horizontally, according to punctuation markers. Short segmentation was preferable where possible.

During comparative analysis of the parallel corpora the deaf expert uncovered several instances of omission or unclear, ambiguous signing of technical phraseology and lexical items related to the semantic field of linguistics. This sparked the realization that there is need for terminological support for interpreters of LIS and that a specialist terminological electronic dictionary could be developed for trainee interpreters of sign language at the Advanced School of Modern Languages for Interpreters and Translators (*Scuola Superiore di Lingue Moderne per Interpreti e Traduttori – SSLMIT),* at the University of Trieste. It was felt that a glossary would serve not only as a terminological tool for students but also as a didactic support to enhance their signing ability by providing signed examples of sentences in context in three languages.

Italian Sign Language paper-based dictionaries tend to be of generic nature aimed at learners of sign language and therefore mainly deal with day-to-day terminology (Angelini *et al.* 1991, Magarotto 1995, Radutzky 1992, Romeo 1991). These dictionaries are essential tools for students learning basic sign language, but of less help to the community or conference interpreter dealing with medical and court terminology, or conferences on topics such as: '*The origins and prehistory of language', 'Video telephony for the Deaf'* or *'Problems of mental health in deaf people'*. Today, computer technology and the widespread adoption of alternative media such as CD-ROMs and DVDs can allow dynamic images of signs to be combined with written information or hypertextual links. Electronic dictionaries of this kind have started to make their appearance in Italy such as *Dizionario mimico gestuale* (Pignotti 1997), *Dizionario Italiano/LIS* (Piccola Cooperativa Sociale "Alba" 2003) and *e-LIS* an online dictionary[7].

Specialist lexis is abundant in English and Italian and hence sign language interpreters are constantly having to cope with it, but face a dilemma: there are few signs in sign languages to translate them. Signed languages everywhere have evolved within a domestic environment and close-knit Deaf communities. The Deaf, in order to communicate, use non-technical everyday language related to family, feelings, food, health, weather, social events and so on. Deaf people may be acquainted with LSP in written form at work, but they rarely need to use it beyond the workplace. Standard signs therefore do not exist in LIS for numerous technical and complex concepts found in spoken Italian or English. Interpreters however, are expected to find a rapid adequate solution and do so by joining together existing signs or inventing new ones. Signed neologisms become established only if transparent enough to convey meaning to the Deaf and if frequently used by other interpreters the same way. If 'technical' signs differ in their configuration from one interpreter to another, this may cause perplexity among the Deaf, as was discovered on analysing the corpus of conference speeches.

4. The glossary

The digital conversion of the corpus into aligned parallel corpora permitted rapid word counts, calculation of word frequencies and the running of concordances of the English, Italian and LIS (glosses) to detect lexical items related to the field in question: linguistics. As the source language of the corpus was English it was decided to manually select specialized terminology from the English word count which initially resulted in around 500 items. Concordances were run for each one using *Wordsmith Tools* and according to the degree of relevance and frequency of use a further selection was made. Thus, the potential pool of specialized English terminology was reduced to approximately 200 items. Rather than compiling a trilingual terminological glossary containing all the items, a pilot version was produced in electronic format on CD-ROM to be tested on students at the SSLMIT[8] (Kellett Bidoli 2005c). Ten lexical items were selected for the pilot version which produced over 60 entries (including synonyms and cross-references) across the three languages. Compilation has continued beyond the pilot version, at present standing at around 100 English entries, and is expected to terminate by mid 2007.

A semasiological approach was chosen leading to an alphabetical ordering of three separate indices: English, Italian and LIS. Headwords and corresponding articles in each of the three languages were first colour-coded, ordered vertically and alphabetically, irrespective of language, before being sent to Turin for transfer onto an HTML application (at the Piccola Cooperativa Sociale "Alba" a r.l – O.n.l.u.s). A 'cross-browser' approach was chosen that will permit access to the completed glossary through a wide choice of browsers and operative systems.

Colour coding of the trilingual entries permitted rapid, visual identification during compilation and will be retained on the CD-ROM and enhanced by two national flags and a LIS label. Below is a monochrome example of the trilingual articles for the lemma **community interpreting.** Where the word <u>IMAGE</u> appears, trainee interpreters will find an icon on which to click in order to obtain a dynamic image of the correctly signed lexical item, or a fully signed version of the example provided below the definition. Trainees will thus be able to obtain trilingual lexical information as well as correct word order sequences and collocations at the click of a mouse.

*****

[7] An on-line dictionary of LIS being compiled at the Istituto di Comunicazione Specialistica e Plurilinguismo, Accademia Europea di Bolzano, Italy (http://elis.eurac.edu).

[8] Kellett Bidoli C.J. (2004). *Glossario inglese - italiano - lingua italiana dei segni (LIS). La lingua dei segni e l'interpretazione: il linguaggio delle conferenze*, (progetto pilota), on CD-ROM, Turin, Piccola Cooperativa Sociale "Alba" a r.l – O.n.l.u.s.

English
community interpreting noun/uncountable
[kəˈmjuːnəti ɪnˈtɜːprɪtɪŋ – ɪnˈtɜːrprɪtɪŋ]
interpretazione in campo sociale [Italian index]
interpretazione in campo sociale [LIS index]
Definition: Interpreting in two language directions consecutively and without notes, principally to assist migrants who cannot speak their host country's language, in order to enable them to gain full access to legal, health, educational and social services.
Example: *Over the past decade the field of community interpreting has increasingly attracted the attention of scholars worldwide.*
Synonyms: public service interpreting, public sector interpreting.
Note: *'Community interpreting' is a form of 'liaison interpreting' which has long been practised but until recently largely ignored as a scholarly subject if compared to 'simultaneous' and 'consecutive' (conference interpreting). There is some debate as to whether 'legal interpreting' and 'court interpreting' can be considered as belonging to 'community interpreting'. 'Sign language interpreting' is considered a form of 'community interpreting' as Deaf people within our societies are often in need of language assistance in social and institutional settings.*
See also: ***liaison interpreting, court interpreting, legal interpreting***

*****

Italian
interpretazione di comunità
community interpreting [English index]
interpretazione di comunità [LIS index]
See: ***interpretazione in campo sociale***

*****

LIS
interpretazione di comunità
interpretazione di comunità [Italian index]
community interpreting [English index]

Definition: Interpretazione bi-direzionale e consecutiva, senza l'ausilio di appunti, con l'obiettivo di assistere principalmente gli immigrati che non parlano la lingua ospitante o i sordi al fine di permettere loro di usufruire dei servizi legali, sanitari e sociali.
IMAGE
Example: *Nell'ultimo decennio il campo dell'interpretazione di comunità ha attirato l'attenzione degli studiosi nel mondo.*
IMAGE
Synonyms: interpretazione in campo sociale
Note: *la '**Interpretazione di comunità' è una forma d'interpretazione di trattativa che si è sempre praticata ma che è stata a lungo ignorata come disciplina di ricerca rispetto alla 'interpretazione di conferenza'. L'interpretazione nella lingua dei segni rientra in questa categoria, in quanto i Sordi hanno spesso bisogno d'interpreti che lavorano in ambito sociale o istituzionale. Nell'italiano esiste il sintagma 'interpretazione in campo sociale' mentre nell'interpretazione vocale si evita di usare 'interpretazione di comunità' in quanto ci si può confondere con l'interpretazione praticata nell'ambito delle istituzioni europee. In LIS invece 'interpretazione di comunità' è il termine più usato.*
See also: ***interpretazione di trattativa, interpretazione in tribunale, interpretazione giuridica***

*****

British English and if necessary American English phonetics are provided followed by colour-coded bilingual translation equivalents of the headword. At first sight there seems little difference between the Italian and LIS equivalents, but by clicking on one or the other, bi-directional access can be obtained to separate articles which have the same definitions and examples but often different notes, with the additional advantage of imagery in the case of the LIS articles. The glossary is tri-directional, so that users can start from an index item or entry article in any of the three languages in order to access information in the other two.

All exemplification of definitions was obtained from concordances of the corpus lexis, run in order to find all occurrences of each item as illustrated below in the extract of a concordance for **chunk** used as a verb or noun:

- not glued to the original. They can **chunk** the information because they
- an interpreter or as a trainee begin to **chunk** the original differently.
- hold more information because each **chunk** will in itself already contai
- rently. You look at larger and larger **chunks** in the original rather than

Concordances led to exemplification in context of each headword, as shown in the following two examples:

*****

English
chunk noun/countable [tʃʌnk]
segmento [Italian index]
segmento [LIS index]
Definition: A portion of discourse of variable length.
Example: *You look at larger and larger chunks in the original rather than just at one particular word.*
Note: *An interpreter trainee should not memorize single words in the source discourse but learn to listen to chunks of information before translating into the target language. The length of a chunk depends on individual mnemonic capacity which expands through training and experience.*

*****

English
chunk verb [tʃʌnk]
segmentare [Italian index]
segmentare [LIS index]
Definition: To mentally select segments of discourse of variable length.
Example: *You as an interpreter or as a trainee begin to chunk the original differently.*

*****

Information is provided on spelling variants and linguistic or semantic features of interest obtained from the patterns of language usage, connotation and colligation that were revealed by concordancing. For example in the case of the head word **communication** three separate observations are noted:

*****

English
communication, noun/uncountable [ˈkəˈmjuːnɪkeɪʃən]
comunicazione [Italian index]
comunicazione [LIS index]
Definition: The ability to transfer thought and feelings beyond the self.

Example: *It's hard to doubt that the early human species had language, not rudimentary gestural communication or grunts and cries, but the ability to make gestures into word signs.*

Note:

- Hearers associate communication with speech, but in parallel their feelings and emotions are conveyed through complex automatic gestures, facial expressions and body movements of which they are often unaware: Non-Verbal Communication (NVC). Signed languages, by excluding speech, emphasize and elaborate upon the NVC.
- frequently related pre-modifiers are: *child -, early -, gestural -, human -, visual -*;
- common collocations: *- begins, - evolves, - of a language, - stage.*

*****

The addition of numerous cross-reference entries (synonyms, related terms, compounds and derivations of interest), not all found in the original corpus, were deemed necessary to provide ample information for trainees and are included as separate entries, often without a complete article, but they guide the user to a headword with a complete one. For example:

*****

English
ear-voice span noun/uncountable
[ɪə vɔɪs spæn – ɪr vɔɪs spæn]
décalage [Italian index]
décalage [LIS index]
See: décalage

*****

The most evident advantage of producing electronic sign language dictionaries is that dynamic images of signs provided by deaf signers versus the static ones of old can be and must be included. Without the inclusion of high quality images the principle aim of visually illustrating sign language is defeated. Terminological data collection and graphic representation take up a major part of the time required to produce sign language glossaries/dictionaries. However, visual representation is equally important. Great care and planning must go into the digital filming with optimal illumination. Details have to be taken into account like the contrast of the signer's clothing with the background and agreement beforehand on how the headwords and examples in LIS should be signed. A major problem during filming remains the translation from voice to sign of technical words and neologisms as well as the fact that single words in English or Italian may not have a corresponding sign at all. Meaning in signed languages is conveyed not only through signs but also through classifiers, non manuals, or fingerspelling. For example the term '**classifier**' used in sign language linguistics literature in English is translated by interpreters into Italian as '*classificatore*' for want of an alternative linguistic term in Italian. '*Classificatore*' in this case, is simply a convenient English loan because in Italian it normally means 'loose-leaf file', or 'filing cabinet' lacking any linguistic connotation. To further complicate matters, in the American literature there is also a distinction between different classifier types: **entity classifiers, handling classifiers, tracing classifiers, quantity classifiers** etc. In LIS no ready made distinction exists in sign; one simply signs CLASSIFIER because research in the field of LIS classifiers is at an early stage and distinctions have not yet been made. If and when identified they may not necessarily fit the American model[9]. In the glossary, in the Italian and LIS indices, these examples and others are translated literally from English into Italian in order to locate them within the indices, but the signer had to find strategies to convey the full meaning without using single, equivalent, readymade signs.

5. Conclusion

Any form of electronic audio-visual support is an invaluable aid for anyone involved in sign language teaching or sign language interpreter training, not only from Italian to LIS, but also from English to LIS in view of the continuing spread of English as an international language of communication. There are numerous advantages in using electronic format in compiling dictionaries of spoken and signed languages singly or in combination:

- the possibility of including dynamic illustration of sign language terminology and its exemplification in context;
- the speed of instant access through hyperlinks to translation equivalents and related terms;
- unlimited space;
- creativity in the form of varied graphics, the use of colour, insets and numerous visual as well as acoustic devices.

Sign language discourse can be filmed, transformed by transcription into glosses and concordanced like any spoken language to study sign patterns and particular usages which should lead to a better understanding of sign language grammar. It is hoped that the methodology described above, which is essentially straightforward and simple (but time consuming), will encourage others to use corpus construction for the collection of samples of authentic discourses containing different LSP genres be they in Italian, or other spoken languages in order to compile terminological dictionaries or specialist glossaries including a signed language.

6. References

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

Caselli, M.C., Corazza, S. (eds.) (1997). *LIS. Studi, esperienze e ricerche sulla Lingua dei Segni in Italia. Atti del 1° Convegno Nazionale sulla Lingua dei Segni*. Tirrenia (Pisa): Edizioni del Cerro.

Gran, L., Kellett Bidoli, C.J., (eds.) (2000). *Signed Language Interpretation and Training: Theoretical and Practical Aspects*. Trieste: EUT.

Kellett Bidoli, C.J. (2004). Intercultural Features of English-to-Italian Sign Language Conference Interpretation: A Preliminary Study for Multimodal Corpus Analysis. In *Intercultural Discourse in Domain-specific English*, *Textus,* 17/1, C. Candlin and M. Gotti eds. (pp. 127-142) Genoa: Tilger.

Kellett Bidoli, C.J. (2005a). Investigation into Linguistic and Cultural Mediation between the English-speaking

[9] Serena Corazza and Luigi Lerose – personal communication.

World and the Italian Deaf Signing Community. In *Cross-cultural Encounters: Linguistic Perspectives*, M. Bondi and N. Maxwell eds. (pp. 159-173) Rome: Officina Edizioni.

Kellett Bidoli, C.J. (2005b). A Multimodal Approach in a Multilingual Context: Interpretation from English to Italian to Italian Sign Language. In *ESP Across Cultures*, vol. 2.

Kellett Bidoli, C.J. (2005c). From voice to gesture: methodological aspects in the compilation of a terminological tri-lingual multimodal dictionary. In *The Interpreters' Newsletter* 12 (pp. 71- 85) Trieste: EUT.

Kellett Bidoli, C.J. (Forthcoming a). The Linguistics Conference Setting: A Comparative Analysis of Intercultural Disparities during English to Italian Sign Language Interpretation. Paper presented at the Seminar on *The Role of English in Institutional Settings: An Intercultural Perspective* at the the ESSE European Society for the Study of English 7th Conference, Zaragoza 8th-12th September, 2004

Kellett Bidoli, C.J. (Forthcoming b). English in the Invisible City. Paper presented at XXII Convegno AIA, *Cityscapes – Islands of the Self*, Cagliari 15th - 17th September 2005.

Magarotto, C. (ed.) (1995). *Vocabolario della lingua gestuale italiana dei sordi*, Rome: Armando Editore.

Ochse, E. (2001). EFL with Adult Deaf Students: Two Cultures, Two Approaches. In *"Languaging" in and Across Human Groups Perspectives on Difference and Asymmetry*, *Textus*, 14/2, G. Cortese and D. Hymes eds. (pp. 447-471) Genoa: Tilger.

Ochse, E. (2004a). Language – English – Difficult//Question – You – Think – What? In *Intercultural Discourse in Domain-specific English*, *Textus*,17/1, C. Candlin and M. Gotti eds. (pp.143-158) Genoa: Tilger, .

Ochse, E. (2004b). A Language via two others: learning English through LIS. In *Representation and Processing of Sign Languages*, O. Streiter and C. Vettori, eds. (pp. 68-74) Lisbon: Centro Cultural de Belem.

Ochse, E. (2005). LIS/Italian as a Basis for Teaching EFL: Cognitive and Intercultural Implications. In *Cross-cultural Encounters: Linguistic Perspectives*, M. Bondi and N. Maxwell eds. (pp. 228-237) Rome: Officina Edizioni.

Piccola Cooperativa Sociale, "Alba"(2003). *Dizionario Italiano/LIS*, on CD-ROM, Turin: Piccola Cooperativa Sociale "Alba" a r.l – O.n.l.u.s.

Pignotti, G. (ed.) (1997). *Dizionario mimico gestuale* (CD-ROM), Ascoli Piceno: Rinascita.

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

Romeo, O. (1991). *Dizionario dei segni. La lingua dei segni in 1400 immagini*, Bologna: Zanichelli.

An Animator of Gestures Applied to the Sign Languages

Rúbia Medianeira Denardi [1], Paulo Fernando Blauth Menezes [1], Antônio Carlos da Rocha Costa [2]

[1] II/UFRGS - Instituto de Informática
Universidade Federal do Rio Grande do Sul
{rmdenardi, blauth}@inf.ufrgs.br

[2] ESIN/UCPel – Escola de Informática
Universidade Católica de Pelotas
rocha@atlas.ucpel.tche.br

Abstract

Motivated for to the expansion of the Internet and the increasing development of Web technologies and for a great number of people with distinct necessities search at that the information they need, we try to attend the deaf community by development an Animator of Gestures applied to the Sign Languages, the AGA-Sign, with the goal of assisting practical writing of signs and in the familiarization with the language. This work presents an application for automatized generation of animations of gestures applied to the Sign Languages from texts written in SignWriting. The used signs for the development of the application had been elaborated from the LIBRAS and the animations had been generated through model AGA (graphical animation based in the Automata Theory).

1. Introduction

The Internet widespread and the increasing development of technologies to the Web join a great number of people with distinct necessities searching the information that they need, using the Internet as a way of education and learning.

The diversity of people who use the Internet as part of education, researches search to develop technologies, methodologies and tools look forward the necessities of these people, particularly the deaf people, for example, (Costa, 2004; Informatics & Telematics Institute, 2004; FIAP, 2004).

Considering the significant number of the special individuals, that according to Census of 2000 of the Brazilian Institute of Geography and Statistics - IBGE, are 5.685.956 special individuals ith educative necessities - PNNE, with hearing difficulties in Brazil, 2002 the National Congress and the Presidency of the Republic had approved and confirmed the Law n.º 10.436, of April 22th, 2002, that recognizes the LIBRAS (Língua Brasileira de Sinais), Brazilian Sign Language, as legal way of communication and expression of the deaf community, and that stimulates the schools to accept as educational resource.

According to Stumpf (2000), the deafness is a difference that makes of its carrying people who pass to communicate of essentially in a visual form, therefore is perfectly a compatible form to the use of the computer as educational tool, which become the equipment and the technology of information processing essencial instruments in the education of deaf people.

In this article presents a model for the automatized generation of animations of gestures, applied to the Sign Languages and the vision of the use of this system, as for example, the aid in the teaching and learning process of deaf people.

We adopted the *SignWriting* system, which can be used for register the written of any Sign Languages, of any country of the world.

From studies carried through on the *SignWriting* and its contribution with the deaf community, considers it generation of animations of signs written in *SignWriting* in Web environment. To generate the animations, we use the animations model based in the Automata Theory, called AGA (Accorsi, 2000), originating the language of AgaML description (AGA Markup Language) (Magalhães, 2002), that structure the animation content in automatons that describe the behavior of synthetic actors during the animation, where each synthetic actor is controlled by a proper automaton. The model for signs animation is called AGA-Sign (Animator of Gestures Applied to the Sign Languages).

The paper is organized as follows. In section 2, we review aspects of the LIBRAS and sumarize the main features of the *SignWriting* system. Section 3 presents the animation model, the AGA-Sign, describing the tools that are part of the model. Section 4 presents the contribution of the AGA-Sign in the teaching and learning processes of deaf people. Section 5 brings the conclusion.

2. LIBRAS and the *SignWriting* system

The Sign Languages used by the deaf people are not universal, each country possess its Sign Languages, with proper grammatical structure and having influences from the native culture (with state and region variations and between specific groups with its slangs). The language condition is attributed to the Sign Languages, and not to an artificial code or simple sets of mimic gestures, because they are composed by the same linguistic levels of the verbal languages: the phonologic, the morphologic, syntactic and the semantics (FENEIS, 2005).

The fact to be used a form of communication and expression of appearance-motor nature, and not of verbal-auditory nature as the verbal languages; do not hinder the Sign Languages to consist in complete linguistic systems for transmission of ideas and facts, concrete or abstract. In Brazil, the Sign Languages is called LIBRAS (Brazilian Sign Languages).

The word equivalent or lexical item, in the verbal-auditory languages is called of sign in the Sign Languages. The signs are formed by way of the combination of forms and movements of the hands and control points in the body or the space.

In the Sign Languages the following parameters can be found that will form the signs:

- Hands Configuration: They are forms of the hands that can be the manual alphabet or other forms made by the predominant hand (right hand for the dexterous or left for the left-handed people), or by the two hands.
- Point of joint: it is a place where the configured predominant hand takes place, or either, place where the sign is made, being able to touch some part of the body or to be in a neutral space.
- Movement: The signs can have a movement or not. The movements can be internal to the hands (movements of the fingers) or displacements of the hands in relation to the body.
- Face and/or corporal expression: Corporal the face expressions/are of basic importance for the real conformity of the sign, being that the tune in Sign Languages is made by the face expression.
- Orientation/Direction: The signs above have a direction in relation to the parameters. Thus, the verbs TO GO and TO COME (Figure 1) if oppose in relation to the direction.

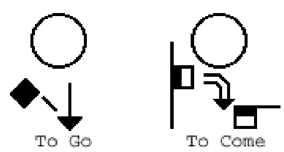


Figure 1: Signs To Go and To Come

The Law N° 10.436, in its article 4° determined and confirmed in 24 of April of 2002 state the following: "the federal educational system and state, community educational systems and of the Federal District must guarantee the inclusion in the formation courses of Special Education, Phonoaudiology and Teaching, in its levels average and superior, of the education of the Brazilian Sign Languages - LIBRAS, as integrant part of the National Curricular Parameters - PCNs, as current law".

About LIBRAS, however, it is not enough to know the signs separately, it is necessary to know the grammatical structure of the phrases of this language, as well as the resources of dialogue.

As cited before, the animation of gestures can be applied to any Sign Languages. However, in this work the signs used belong to the LIBRAS, for being the native language of the Brazilian deaf people.

In this paper, the signs were writing with the *SignWriting* symbols.

The *SignWriting* was created by Valerie Sutton founder of DAC (Deaf Action Committee), a system of writing for Sign Languages (Sutton, 1990).

It was developed to be a form of written for Sign Languages, as well as the diverse alphabetical, syllables notations and ideographic of written forms of verbal languages. An evident and decisive difference, however, is that the last ones had a development of historical character, to the step that *SignWriting* rationally was conceived, being, therefore a formal language. The graphical expressions of the *SignWriting* (Figure 2) restrict to describe movements (physical), as well as face expressions, and not meant it of the signs, making with that the system can, therefore, to represent any Sign Languages.

Further than the graphical character, the *SignWriting* was conceived to be registered in fixed way, in paper. This reflects inside in its more than enough repertoire of dedicated symbols to the representation of the dynamics of the corporal elements in game of the signs (Costa, 1997).

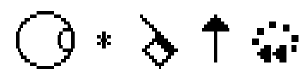


Figure 2: Example of the SignWriting symbols

The *SignWriting* is used in more than thirty countries and in Brazil it started to be used in 1996. It had been formed work groups with no deaf people and deficient in hearing for the diffusion of the system and education of the reading and the writing of the deaf community. The research groups had also been formed to assist in the development of tools that make use of the system, between which it is in case the described system in the present work.

The application of the *SignWriting* to the LIBRAS producing the Brazilian sign writing supplies as adequate tool so that the deaf students fulfill the objective to register for writing its visual language. The productions of signs are made through editors who display *SignWriting* symbols.

3. AGA-Sign

In this section, we present all the tolls that integrate the AGA-Sign, as Figure 3.

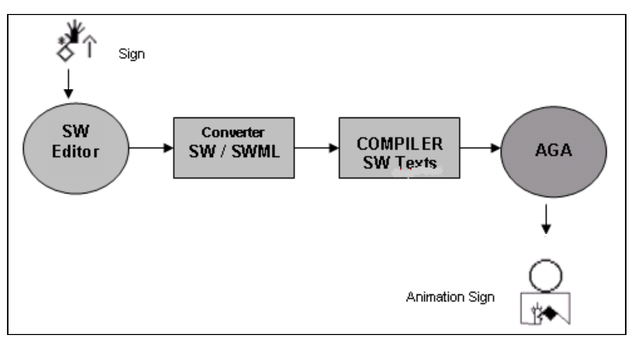


Figure 3: AGA-Sign Model

3.1. *SignWriting* Editors

The signs are represented through editors who process signs in visual way with *SignWriting* symbols, making possible the users of the Sign Languages to be able to write texts in its native language. They make use of symbols of movements, face format of hands, expressions and still a Dictionary of Sign Languages, where the users can store signs, copy them and stick them in documents while they type.

The SignWriter program (Sutton et. al., 1995), the first computer editor for sign languages, defined such an encoding for *SignWriting*.

In this work the signs had been produced in SW-Edit Editor (Torchelsen et al 2002) whose one of the differentials in comparison with others publishers are that it makes possible that the archives are safe in SWML (presented in the next subsection).

3.2. Converter SW/SWML

To suit possible the use of the *SignWriting* system in WWW pages a converter SW/SWML was developed. SWML (Costa, 2005) is a proposal for a general encoding format for *SignWriting* documents, using XML-based markup language for Sign Language Processing, for the storage, changes and processing of texts of the *SignWriting*. With the SWML the document interchange is possible between different programs, the independent analysis of texts of the publisher and also it serves as a format of storage of texts.

After the edition of the signs, the archives of resultant signs are converted for texts SWML, generated for Converter SW/SWML. A text converted into SWML presents the position, rotation, variation, fill and shape of each symbol, as it shows the Example 1.

Example 1 *The SWML representation of LIBRAS sign for* To Show *(written as in Figure 4) is:*

```
<signbox>
   <symb x="53” y="57" x-flop="0" y-flop="0" color="0,0,0">
     <category>01</category>
     <group>05</group>
     <symbnum>001</symbnum>
     <variation>01</variation>
     <fill>03</fill>
     <rotation>01</rotation>
   </symb>
   <symb x="50" y="79" x-flop="0" y-flop="1" color="0,0,0">
     <category>01</category>
     <group>01</group>
     <symbnum>001</symbnum>
     <variation>01</variation>
     <fill>01</fill>
     <rotation>02</rotation>
   </symb>
   <symb x="74" y="69" x-flop="0" y-flop="0" color="0,0,0">
     <category>02</category>
     <group>05</group>
     <symbnum>001</symbnum>
     <variation>02</variation>
     <fill>03</fill>
     <rotation>01</rotation>
   </symb>
   <symb x="30" y="64" x-flop="0" y-flop="0" color="0,0,0">
     <category>02</category>
     <group>01</group>
     <symbnum>001</symbnum>
     <variation>01</variation>
     <fill>01</fill>
     <rotation>01</rotation>
   </symb>
</signbox>
```

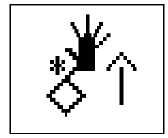


Figure 4: Sign To Show

3.3. Compiler of *SignWriting* Texts

In this section presents an important stage of this work, the detailed study of the *SignWriting* symbols for to know its characteristics and its meanings.

The *SignWriting* symbols are classified as its categories and follow a specific order, called of the SSS (Sign-Symbol-Sequence) (Table 1) and represented for

SSS = (C, G, S, V, F, R)

and represents respectively: *Category*, *Groups*, *Symbol Number*, *Variation*, *Fill* and *Rotation*.

Symbol	C	G	S	V	F	R
*	02	01	001	01	01	01

Table 1: SSS of a SignWriting symbol

Actually, the *SignWriting* symbols are organized in ten categories, as Figure 5.

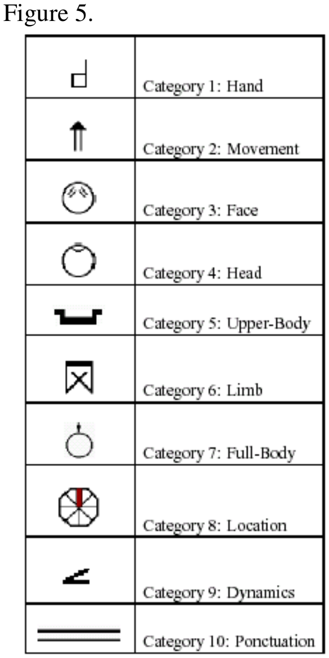


Figure 5: SignWriting Categories

The texts written in *SignWriting* are storages in archives of signs. To visualize the SSS of the symbols that compose a sign, we use the Converter SW/SWML. The SWML documents list all the information of the sign writing in *SignWriting*, suiting possible the analysis of the symbols that are parts of sign. From the knowledge of the symbols and its characteristics, mainly of the movements, it is possible to determine some rules that will be responsible for the animation of the signs.

The movement symbols can represent displacements: rectilinear vertical or horizontal lines; circular horizontal or vertical; and arched horizontal or vertical (Figure 6).

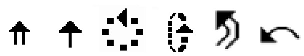


Figure 6: Example of the movement symbols

The movement symbols can be modified to indicate particular aspects of a represented movement, as the greater or minor covered distance (Figure 7), or the direction of the movement (Figure 8).

Figure 7: Example of the variations of vertical movements

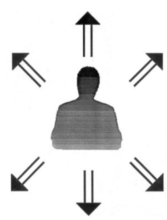


Figure 8: Example of the rotations of vertical movements

For example, the sign presents in Figure 9 have a short displacement of the hands, represented for the vertical movement symbol with *variation* 1.

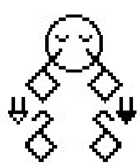


Figure 9: Writing of the sign To Cry

The hands displacement is represented in the Figure 10, where P_i is the initial position and P_f is the final position of the hands.

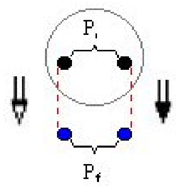


Figure 10: Displacement for the sign To Cry

Then, from variation and rotation is possible to know the direction and the displacement of the object during the animation of the sign. The rule that define this movement is represented in Example 2.

Example 2 *Algoritmic example of the rules for vertical movements.*

```
if (movement = vertical) then
    if (variation = short movement) then
        d ← distance equivalent to the short displacement
        if (rotation = arrow for low) then
          displacement of the object ← (xi, yi - d)
        end if
    end if
end if
```

The displacement of the object happens through of the initial position of the object (x_i, y_i) increased or diminished of the equivalent distance to movement (short, medium or long). In Example 2, the displacement of the object is for low, on the vertical axis (y), not modifying the position on the horizontal axis (x).

The goal of the Compiler of *SignWriting* Texts is the generation of a document that serves as input archive for AGA animator. From the reading of a document SWML the symbols are interpreted and, as the rules, the Compiler generates an AgaML document. This document presents the actors specification, tapes and instances for the AGA actors. The AgaML and AGA model will be presents in next subsection.

3.4. AGA

The model of animations for Web, called AGA (Graphic Animation based on Finite Automata) (Accorsi, 2000) based on the Theory of the Automatons is used to generate the animations. The AGA specifies the animation from a set of actors (objects) and its respective variations during the animation. The specifications in AGA are supported by a formal model based in automatons with exit (Hopcroft et al 2000; Menezes, 2005).

In the AGA, the animation actors are specified through an extension application for the automaton with exit, which attach the variations in the graphical form of the actor from the output of the automaton. In this way, when the automatons are simulated, by means of the reading of an input tape, the transitions between its states control the actor's animation (Accorsi, 2000). In the Sign Languages animation each symbol represents an actor, as Figure 11 and Figure 12.

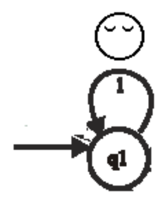


Figure 11: Head Actor specified in AGA.

The Figure 11 and 12 illustrates the basic structure of model AGA specification applied to an actor animation. The symbols are of the To Cry Sign (Figure 9). The actors are specified from automatons with exit where graphical representations are associates to the transitions. These representations correspond to the graphical variations that the actor can suffer during the animation. Of this form, when the hand-right actor carries through a transition of state 1 for state 2, the graphical representation of the actor presented in the animation is modified by the hand-right with bended finger.

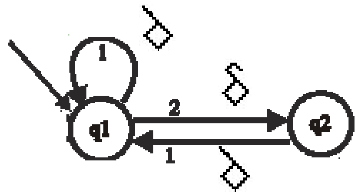


Figure 12: Hand-right Actor specified in AGA

The choice for the AGA was given due the characteristics that facilitate its application in the specification and control of animations in the Web, which can be cited: storage space has supported the recovery of information and maintenance of the content of the animations. From the AGA it originated language of description of the called animation AgaML (AGA Markup Language).

The AgaML organizes the specification of the animation from three basic components: the specification of AGA actors (Example 3), the specification of input tapes (Example 4) and the creation of the instances of the actors (Example 5). The instances can be understood as the association of the specification of an actor AGA with an input tape. The specifications of the actors can be used by some different instances, as well as, can be shared by diverse animations if be stored in independent archives.

Example 3: *Element ACTOR for the specification of the hand-right actor (Figure 12)***.**

```
<ACTOR ID = “hand-rightact” TYPE=”GRAPHICS”
STATES=”2” SYMBOLS=”2”>
    <OUTPUT ID=”1” SOURCE=”01-01-001-01-01-02.gif”
    x=”180” y=”160”/>
    <OUTPUT ID=”2” SOURCE=”01-01-007-01-01-02.gif”
    x=”180” y=”155”/>
    <DESCF>
        <DESCRIPTION STATE=”1”> Strained
    finger</DESCRIPTION>
        <DESCRIPTION STATE=”2”> Bended
    finger</DESCRIPTION>
    <DESCF>
    <TRANSF>
        <FROM STATE=”1”>
                <TO STATE=”1” SYMBOL=”1”
          OUTPUT=”1”/>
                <TO STATE=”2” SYMBOL=”2”
          OUTPUT=”2”/>
        </FROM>
        <FROM STATE=”2”>
                <TO STATE=”1” SYMBOL=”1”
    OUTPUT=”1”/>
        </FROM>
    </TRANSF>
</ACTOR>
```

Example 4: *Element TAPE for the specification of the input tape.*

```
<TAPE ID = “hand-righttape”>
    <CEL SYMBOL=”1” TIME=”50”/>
    <CEL SYMBOL=”2” TIME=”200”/>
    <CEL SYMBOL=”1” TIME=”50”/>
</TAPE>
```

Example 5: *Element INSTANCE for the creation for instances.*

```
<INSTANCE ID = “hand-rigth” ACTOR=”hand-rightact”
ORDER=”2”>
    <USE TAPE= “hand-righttape”/>
</INSTANCE>
```

The program of visualization, AGA Player, was developed in JAVA in format applet and it is executed in the client to realize the reproduction of the animation, as the specifications in AgaML. Some examples can be seen in http://www.inf.ufrgs.br/~rmdenardi/aga/animacao.html.

4. AGA-Sign assisting in the education of Sign Languages

In Brazil, the LIBRAS started to be used in nineties. Because was the question of the Portuguese language, in relation the oral communication, be a very slow process in the communication of the deaf people. It was verified that the LIBRAS are a facilitator not only by the communication, but also of the dissemination of the information (Public Education, 2004).

Stumpf (2000), telling its experience on the use of the *SignWriting* in the Special School Concord, affirms that many deaf pupils when they learn to write think that the written Portuguese is the representation of the Sign Languages that they use. When the pupils start to learn the writing of signs they obtain to separate and to see that it is another language. The two languages working separately and comparing them the result it will be better because he is thus that the learning of one second language happens.

The deaf people can produce excellent materials in its written manifestations as: literature, poetry and texts, if possess the necessary control of the instrument. For the common, this does not happen in the Portuguese language because the learning difficulty of the verbal language for the deaf people is enormous.

For this reason, there is a great interest of the deaf people in learning the Sign Languages and, more recently, in using the *SignWriting* system.

The tools that become possible in the specification of the AGA-Sign, and the appropriate animator, contribute with the learning of the Sign Languages. In that the student says follow to the writing for the fact to be tools in which can produce the signs and store them in dictionaries of signs (editors) and to make animation them (AGA-Sign), assisting in practical of the writing of signs and the familiarization with the language.

The learner of the Sign Languages will also be able to use the animator as a verifier of correction of signs: in the doubt if the written sign he is or not what he does want to represent, the user makes animation the sign to have the confirmation.

5. Conclusion

The AGA-Sign, with its specification for signs animations through *SignWriting* symbols and the tools cited in the work, can be great allied in the teaching and learning of Sign Languages.

The generation of animations of signs through AGA model can contribute with the advance of the research in this area and for the insertion of the deaf people in the world of the technologies of the information, especially the Internet.

The application of the *SignWriting* to the LIBRAS producing the Brazilian writing of signs is an adequate tool so that the deaf students fulfill the objective to register for writing its visual language.

6. References

Accorsi, F. (2000). Animação Bidimensional para World Wide Web Baseada em Autômatos Finitos. *Proceedings of WMF'2000 - 3rd Workshop on Formal Methods*. (pp. 122 – 127)João Pessoa, Paraíba.

Costa, A. C. R (1997) Aplicações de Sign Writing. http://gmc.ucpel.tche.br/signwriting/SWProj.html (Published 1997; accessed Sep 2000)

Costa, A. C. R (2004) Site Projeto SignNet. http://sign-net.ucpel.tche.br/ (Last updated 2004; accessed Out 2004)

Costa, A. C. R (2005) SWML - *SignWriting* Markup Language. http://swml.ucpel.tche.br (Last Updated 2005; accessed Mai 2005)

FENEIS, (2005). LIBRAS. http://www.feneis.com.br (Last Updated 2005; accessed Out 2003)

FIAP (2004) Projeto De olho na *WEB*! http://www.fiap.com.br/portal/int_cda_conteudo.jsp?ID=9647 (Last Updated 15 Mai 2005; accessed Out 2004)

Hopcroft, J. E.; Ullman, J. D. (2000). *Introduction to Automata Theory, Languages and Computation*. Addison Wesley Publishers, 2nd. Edition. Boston, USA.

Humanoid Animation Working Group. (2004). HANIM. http://www.h-anim.org/ (Last update 2004; accessed Abr 2004)

INFORMATICS & TELEMATICS INSTITUTE (2004) Vsigns. http://vsigns.iti.gr (Published 2004; accessad Feb 2004)

Magalhães, G. C. (2002). Aga Player: Animação 2D Baseada em Autômatos para a Web. *Trabalho de Conclusão de Curso.* 41p. *Instituto de Informática, UFRGS*. Porto Alegre, RS.

Menezes, P. B. (2005). *Linguagens Formais e Autômatos*. Editor Sagra Luzatto, Porto Alegre, Brasil.

Public Education (2004). Educação em Libras: há 145 anos, o Ines promove educação para surdos. http://www.educacaopublica.rj.gov.br/jornal/materia.asp?seq=200 (Published 21 Sep 2004; accessed Apr 2005)

Stumpf, M. R. (2000). Língua de Sinais: escrita dos surdos na Internet. *Proceedings of RIBIE - V Congresso Iberoamericano de Informática Educativa (*pp 4 - 6).Viña del Mar, Chile.

Sutton, V. (1990). Sign Writing. http://www.signwrting.org (Last Updated 2005; accessed Mai 2005)

Sutton, V.; Gleaves, R. (1995). *SignWriter – The world´s first sign language processor*. Center for Sutton Movement Writing, La Jolla.

Torchelsen, R. P.; Costa, A. C. R.; Dimuro, G. P. Editor para Textos Escritos em *SignWriting*. *Proceedings of XVIII Conferencia Latinoamericana de Informática.* (v. 1. pp. 1-10) Montevideo, Uruguai.

Processing linguistic data for GSL structure representation

Eleni Efthimiou[1], Stavroula-Evita Fotinea[1], Galini Sapountzaki[2]

[1] Institute for Language and Speech Processing
[2] Greek Sign Language Lab, University of Thessaly
eleni_e@ilsp.gr, evita@ilsp.gr, gsapountz@uth.gr

Abstract

The here presented work reports on incorporation of a core grammar of Greek Sign Language (GSL) into a Greek to GSL conversion tool. The output of conversion feeds a signing avatar, enabling dynamic sign synthesis. Efficient conversion is of significant importance in order to support access to e-content by the Greek deaf community, given that the conversion tool may well be integrated into various applications, which require linguistic knowledge. The converter is built upon standard principles of Machine Translation (MT) and matches Greek parsed input to equivalent GSL output. The transfer module makes use of NLP techniques to enrich linear sign concatenation with GSL-specific complex features uttered both manually and non-manually. GSL features are either checked against properties coded in a lexicon DB for base signs or they are generated by grammar rules. The GSL computational grammar is based on natural data analysis in order to capture the generative characteristics of the language. The conversion grammar of the transfer module, however, makes use of a number of heuristic solutions. This is implicated by the type of input for conversion, which derives from a statistical shallow parser, so that various semantic features have to be retrieved by mere grouping of lemmata. However, this type of input is directly connected with the requirement for fast processing of vast amounts of linguistic information.

1. Introduction

GSL sign synthesis originally involved dynamic generation of single signs (word level linguistic units). In this framework, a library of sign notation features has been converted to motion parameters of a signing avatar (Karpouzis et al., 2005). These features, allow to represent the “phonological structure” of any sign and along with a set of GSL specific features relevant for sign formation, accompany sign lemmata in a multipurpose lexicon data base (Efthimiou et al., 2004). Exploitation of sign synthesis to access e-content, required to extend synthesis to phrase level. A computational grammar based on Unification Grammar principles (Shieber, 1992) is developed to provide for generation of GSL structures.

For the representation of the phonological features of GSL the extended HamNoSys annotation system (Prillwitz et al., 1989; Hanke, 2002) has been adopted. Sign coding is further enriched to provide for the non-manual obligatory features, which accompany hand action in order to make linguistic content fully acceptable. Mouthing patterns, facial expressions and body/shoulder movement -also used for the indication of phonetically (stress) or syntactically uttered (focus position in sentence) elements of the linguistic message in spoken languages- comprise the multi-layer information coded in the GSL lexicon DB. Eyebrows movement and eye gaze are also coded, when present, since they are significant obligatory parts of GSL sign formation.

The computational grammar GSL can handle sign phrase generation as regards the basic predicate categories and their complements, and extended nominal formations. The rules generate surface structures with a linear ordering that corresponds to basic sign sequences in a phrase. Maximal phrase feature bundles (Carpenter, 1992) contain features for both linear and non-linear linguistic information.

Here, we report on how these resources are exploited in the environment of a conversion tool that matches structured chunks of written Greek to GSL structures. Emphasis is put on structure matching between the two languages, and coverage of grammar phenomena of GSL.

2. Greek to GSL converter description

The converter (Fotinea et al., 2005) is programmed in Java to allow for quick and efficient design development compatible with all system platforms. XML technology has been utilized as a means for describing structured documents in a re-usable format, while Java technology contains embedded tools for the management of XML texts. Hence, the converter utilizes multi-level XML-based annotated sentences, exploiting XML technologies for its collaboration with the shallow parsing sub-module that creates the Greek parsed chunk input for conversion.

The conversion tool performs top-down, rule-based meta-syntactic analysis. Rules are organized in three sets, the structure set, the chunk set and the feature set. The structure set allows for linguistic actions involving (conditional) re-ordering of chunk sequences to reflect the morpheme order of GSL. A second set of rules performs on the chunk level, allowing for addition, deletion or modification of specific entries, whereas a third set of rules applies to feature level, to perform either insertion of mostly GSL specific features, or modification or deletion of existing features, if required for GSL synthesis. Provision has been made that the user may arrange rules into user-defined rule sets, allowing for execution/testing of either all rules or any given combination. Rule execution is iterative and for each iteration all rules are examined, the output of each rule serving as the input of the next, provided that the rule context (‘if-part’) is satisfied. Iterative execution continues as long as change of the input occurs. In Figure 1, a screen shot of the application environment is depicted. The upper half of the screen shows chunked input and the bottom half depicts the output (single rule execution example).

3. GSL grammar coverage

The computational grammar currently handles analysis and generation of both clause- and phrase-level phenomena of GSL, which demand both linear and multi-layer handling. Structures are enriched with GSL-specific features related to the various phenomena of the language.

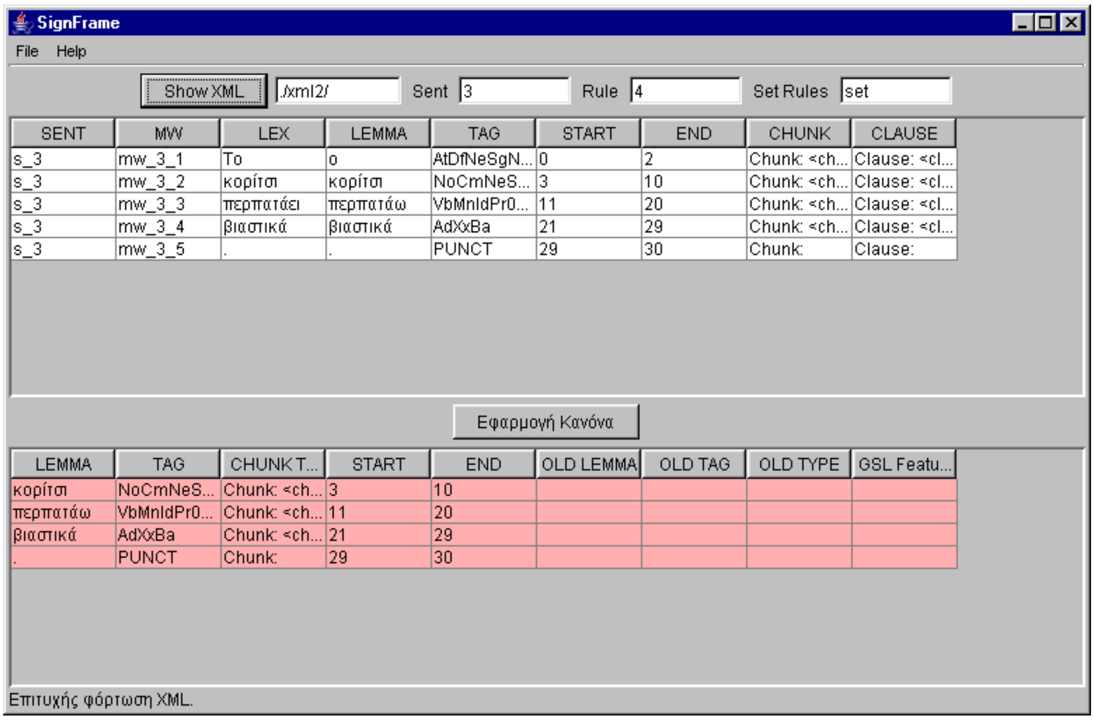


Figure 1: The application environment of the conversion tool.

A subset of these features is coded as lemma-related properties in the GSL lexicon database, and they acquire specific values in rule descriptions after search in the lexicon. A typical example is plural formation where NP plural value resulting from agreement checking inside NP, receives the GSL-specific feature for plural morphology that is coded to the lemma of the base sign (head of the construction). This type of morphological enrichment is required in order to allow for correct representation of the analysed phrase by an avatar in a computational environment (see also on the same subject Marshall & Safar (2005)), but also reflects the morpho-syntactic rule of the grammar a human signer utilizes when uttering the specific phrase.

The lexicon that interacts with the computational grammar codes both articulatory and morpho-syntactic features of lemmata. Figure 2 depicts part of the fields of grammar information coded in the lexicon database, where 2(a) shows the manual and non-manual obligatory features for sign articulation, and 2(b) depicts indicative morpho-syntactic features related to lemma entries.

As regards lemma formation (Figure 2(a)), the 'yes' value in the field for eye gaze as well as the different mouthing values display obligatory simultaneous performance with HamNoSys annotated hand motion. For example, plural 'YOU' (id 26) as coded in the lexical DB, demands obligatory eye gaze performance (towards the addressee).

Morpho-phonological properties of GSL, as for example plural properties (Figure 2(b)) are coded in the field 'np_plural', the different values of which correspond to plural formation with numeric value or quantifier on singular sign (i.e. '2, 3, … days', '2, 3, … pencils' etc), to formation with movement repetition and/or change in space (i.e. as in the case of 'book', 'tree' and 'child') and to 2handed formation, if the singular sign is formed with one hand and not body anchored (i.e. 'airplane').

The semantic values related to the field 'GSL_aspect' provide information on language intrinsic adverbial properties for the definition of continuation, duration, degrading, intensity or repetition related to the action indicated by the predicate. 'GSL_aspect' value 'dur' indicates that the sign movement continues for longer than default, 'dim' signifies small span of movement to indicate minimal action/event (i.e. with predicative base signs such as 'wind-is-blowing', 'I-walk', 'I-speak', 'I-eat' etc), 'int' denotes bigger span and abrupt pauses in movement (i.e. with signs as 'feel-a-pain', 'it-rains' etc) and 'rep' indicates obligatory repetition of sign movement with interval pauses (i.e. with signs as 'ask' or 'travel').

A number of parameters related to traditional linguistic analysis have been taken into account, in order to decide on the structures to be adopted as the default output of analysis. As regards predicate classification, empirical evidence and related analysis (Sapountzaki, 2005) support three main clusters 'Simple Predicates', 'Predicates of Direction' (i.e. predicate 'give') and 'Spatial Predicates' (predicates of movement in real space, i.e. 'walk'). The current grammar implements a pattern which incorporated both simple and spatial predicate formations. Predicates of direction are not yet treated, since they heavily involve use of classifiers that are not yet implemented in the conversion grammar.

As regards word order options, two orders for clause-formation appear to be acceptable in a wide range of data related to the predicate categories under consideration. The one involves strings of the type [Agent-Predicate-Complement], whereas the other arranges constituents in [Agent-Complement-Predicate] strings.

Id	Lemma	HamNoSys	Mouthing	Eye gaze
4	ΤΡΕΧΩ	[illegible]	OL7	
13	ΜΑΛΩΝΩ	[illegible]	CN17	
5	ΚΑΤΗΓΟΡΩ	[illegible]	CN17	
3	ΦΙΛΩ	[illegible]	CN14	
26	ΕΣΕΙΣ	[illegible]		YES

(a) Manual and non-manual obligatory features for sign articulation.

Id	Lemma	Word family	GSL aspect	Np plural	Can be a topic	Noun verb modification	Syntactic movement	Real movement	Becomes classifier	Combines with classifier
13	ΜΑΛΩΝΩ	επιθυμία φωνάζω			No	No	No	No		No
5	ΚΑΤΗΓΟΡΩ	επικίνδυνος			No	No	No	No		No
2	ΕΒΔΟΜΑΔΑ		Rep	1	Yes					
31	ΑΓΑΠΩ	αισθήματα	Dim/Int		No	Yes	No	No		No
7	ΑΓΟΡΙ		Dim/Int	1	Yes					Yes

(b) Grammar features related to lemma entries.

Figure 2: Fields of grammar information coded in the lexicon database.

For reasons of computational efficiency, implementation has adopted the [Agent-Complement-Predicate] arrangement, given that the specific order is supported by theoretical analysis (Efthimiou, 2006) as the basic word order of the language and also allows for an adequate handling of the set of phenomena that take place on clause level (sentential negation, tense declaration, interrogation, etc). Adoption of this order also facilitates handling emphasis assigned to either predicate arguments or various sentential adjuncts (i.e. temporal adverbs).

Surface deviations of the acknowledged concatenation order as regards main constituents of the clause are treated as cases of emphatic structures. Our approach provides for a clause-initial position undefined for grammatical category, which serves as a place-holder for emphasis, similar to an analysis proposed for the Greek language (Efthimiou & Zombolou, 1995). In the cases where this position is filled, the surface linguistic data seem to deviate from the standard constituent concatenation patterns of the language.

This however, is not true, if we adopt the structure pattern for clause formation: [Emphasis_Position-Agent-Complement-Predicate]. When no emphasized constituent is present, the clause-initial position remains void. Otherwise, any constituent may fill the clause-initial position, receiving accordingly the interpretation of emphasis. In this sense, the Emphasis_Position is free for any semantic category, including the Agent, the Patient, the Beneficiary as well as all types of phrases with adverbial value (locative, temporal, etc).

Various operations inside the NP, mainly involve constituent arrangement around the head, including actions of deletion of information irrelevant to articulation in 3D space (i.e. determiner deletion), and feature insertion obligatory for the reconstruction of information articulated in a multi-layer manner in GSL (i.e. mouthing patterns parallel to head sign formation for quantitative adjectives). Special provision is made that when the head of the input NP is characterized as proper, instead of searching the bilingual lexicon, a finger spelling procedure is activated for the representation of the string of characters forming the proper name.

A detailed description of the phenomena that currently comprise the GSL computational grammar follows, along with a discussion on handling the matching parameters implemented for the needs of the conversion operation.

4. Rule description

4.1. Clause level operations

4.1.1. Sentence word order

As already stated the default clause formation order takes into account concatenation instantiations of one- and two-place predicates along with options for various sentential adjuncts. Whereas for main clause constituent arrangement, the predicate systematically fills the string-final position, in the case of two-place predicates the Agent always precedes the predicate Complement, resulting in strings as in the examples below (Ex.1-2).

(Ex.1) BOY COME
= A boy comes

(Ex.2) I TEACHER LIKE
= I like the teacher

Temporal phrases are placed in clause-initial position, reserved for emphasized constituents (Ex.3).

(Ex.3) FRIDAY I CHURCH GO
= On Friday I go to the church

However, Greek temporal adverbs such as 'χθες (=yesterday)', 'αύριο (=tomorrow)', 'σήμερα (=today)' are treated by special lexical rules that obligatorily delete the Greek lemma and incorporate the temporal value of the adverb as a complex eye-gaze and head-movement feature on the GSL predicate. An example, shown next (Ex.4), presents the output of the conversion operation after having applied clause structuring, deletion of input lemma and feature insertion for multi-layer representation.

(Ex.4) I CHURCH GO + EYE_GAZE + HEAD_MOVEMENT(RECENT_PAST)
= *Yesterday I went to the church*

The conversion operation related to Ex.4 is sketched in Figure 3, where the left-hand part indicates written Greek chunked input and the right-hand part the resulting GSL structure. The GSL recent past feature (‘GSL_Rec_Pa’) on the predicate lemma indicates activation of obligatory eye gaze and related head movement.

4.1.2. Sentential negation

Sentential negation is treated as required by the adopted clause concatenation order, that is, in the output of the conversion operation a negative particle is always adjuncted to the clausal predicate as discussed in Section 4.2 (Verb Group operations). As regards theoretical analysis see also Antzakas & Woll (2002).

4.1.3. Existential verb deletion

GSL does not make use of existential predicates, like ‘be’. In order to convert Greek sentences into GSL the existential verb has to be deleted and a pause has to be inserted between Agent and Attribute, where the tense indication (except present tense) has to be transferred to the output and be represented with a temporal sign. For the example below (Ex.5), the converter rule for existential verb deletion is given in Figure 4.

(Ex.5) YANNIS+FING_SPELL+PAUSE DOCTOR
= *John is a doctor*

4.1.4. Deictic subject doubling

If the input string contains a pronominal element characterized as ‘strong’ in the Greek analysis notation (opposite to ‘weak’ that corresponds to clitic pronouns of Greek), the right-hand side object has to contain the GSL equivalent to the full personal pronoun, which in this case is the deictic pronoun. Deictic Agent, if present, has to be repeated at the end of the utterance (mainly for verification of Agent information in the case of lengthy utterances). An example is given in Figure 5, which results in strings as in examples (Ex.6-Ex.7).

(Ex.6) HE+DEICTIC COME HE +DEICTIC
= *He comes*

(Ex.7) I+DEICTIC BOOK WANT I+DEICTIC
= *I want the book*

4.2. Verb Group operations

The predicates currently treated in the grammar, present a number of characteristics (Fischer, 1996) which differentiate them from the predicates of direction. In the clause output, the arguments of simple predicates are uttered as separate signs, following the concatenation order(s) of the language. As concerns the predicate articulation, location and direction of movement remain constant, whereas sentential negation, when present, is realized with the utterance of a negative particle strictly following predicate articulation (Ex.7-8).

(Ex.7) I TEACHER THIS LIKE
= *I like this teacher*

(Ex.8) ________ neg
I TEACHER THIS, LIKE NOT
= *I don’t like this teacher*

An exception to general negation rule present predicates which allow for the expression of negation by morphological means inside the base sign. In this case, negation of the semantic content of the predicate is realized by applying mirror image movement (reverse movement) as to start position, i.e. in the case of negation of the predicate ‘want’ (Ex.9). For such predicates it is necessary to treat sentential negation by a lexical rule that matches the input lemma plus negation features, with a separate sign lemma.

(Ex.9) _______ neg
I BALL THIS NOT-WANT
= *I don’t want this ball*

A feature insertion operation involves incorporation of adverbial semantic values in the predicate morphology. To treat elements as i.e. ‘much’, ‘a little’, ‘continuously’ -expressed in GSL on predicate morphology- the current implementation, activates list searching of adverb - predicate cluster combinations in the input string, in order to assign a specific value to the predicate feature ‘GSL_aspect’. Indicative examples are presented next (Ex.10-11).

(Ex.10) I EAT+GSL_ASPECT=dur I
= *I eat a lot*

(Ex.11) RAIN+GSL_ASPECT=int
= *It rains heavily*

4.3. Noun Phrase operations

NP formation in GSL typically lacks open determiner declaration where a number of specifiers, such as qualitative adjectives, are incorporated in base sign articulation as extra (mouthing) features. The lexicon codes base sign articulation as to manual and non-manual obligatory parameters. Any context-dependent information on the base sign has to be reconstructed by rule-based feature insertion.

4.3.1. Article deletion

If the analysis of Greek input string has recognized the existence of a determiner inside NP, then a deletion operation is performed. In the examples 2 and 3 (Ex.2-3) above, one can see the result of this operation.

4.3.2. Adjective absorption

Adjectives are either listed in a concatenation of separate signs, adjuncting properties on the head sign, or they convey their semantic properties, by being uttered simultaneously with the head sign as additional multi-layer features.

In the latter case, they are uttered as a (combination of) facial expressions, simultaneously performed with base sign articulation. A typical instantiation of the above involves expression of qualitative adjective values like ‘nice/good/ugly’ etc. These values correspond to different mouth patterns in GSL. To resolve this type of conversion problem from Greek, a similar approach has been adopted as the one applied for addition of adverbial values to predicates. In this case, the list of lemmata to be translated to features includes the clusters of different adjectives.

Example (Ex.12) illustrates the multi-layer structured NP “nice apple”, being the output of application of the relevant lexical rule, for the case of qualitative adjectives.

(Ex.12) APPLE+MP_CN17
= *nice apple*

```
(SENT   <S>
        SYN     [cl
        SYN     [advp
        TOK     χθες      χθες      AdXx    ad_temp
        SYN     /advp]
        SYN     [np_nm
        SYN     *sing
        TOK     Ένας      ένας      AtIdMaSgNm
                atidsgnm
        TOK     άνθρωπος            άνθρωπος
                NoCmMaSgNm          nosgnm
        SYN     *sing
        SYN     /np_nm]
        SYN     [vg
        SYN     *sing
        TOK     έφαγε     τρώγω
        VbMnIdPa03SgXxIpAvXx                vb_sg
        SYN     *sing
        SYN     /vg]
        SYN     [np_ac
        TOK     τα        ο         AtDfNePlAc
                atdfplac
        TOK     μήλα      μήλο      NoCmNePlAc
                noplac
        SYN     /np_ac]
        SYN     /cl]
        PTERM_P           .         .
        PTERM_P                     punct_fs
        )SENT   </S>
```

```
(SENT   <S>
        SYN     [cl
        SYN     [Agent
        SYN     *sing
        LEM     άνθρωπος
        SYN     *sing
        SYN     /Agent]
        SYN     [Compl
        SYN     *plural
        LEM     μήλο     GSL_Pl01
        SYN     *plural
        SYN     /Compl]
        SYN     [Pred
        SYN     *sing
        LEM     τρώγω
        Id03Sg+GSL_Rec_Pa
        SYN     *sing
        SYN     /Pred]
        SYN     /cl]
      )SENT   </S>
```

Figure 3: Multi-layer indication of recent past in GSL.

```
RULE                                    % Existential verb deletion rule
If Clause=*
     [np_nm][vb_eimai_id][np_nm]
    *
THEN Clause=*                           % delete existential verb
    [np_Agent]= [np_nm]
    [np_Attribute]= [np_nm]
    GSL_Tense=Read_Tense_from_Predicate_Attributes(vb_eimai_id);
    *
```

Figure 4: Existential verb deletion rule.

```
RULE                                    % Deictic pronoun subject doubling when pronoun is present
If Clause=*
     [np_nm and "Pn*St"] % Εγώ, εσύ, αυτός... Pronoun*strong
     [vg] % verb group
    *
THEN Clause= *
        [np_Agent_Deictic]=  [np_nm]
         *
        [vg]
        [np_Agent_Deictic]              % deictic pronoun subject doubling
```

Figure 5: Deictic pronoun subject doubling when pronoun is present.

4.3.3. Adjective concatenation

In general, adjectives, which are not represented by a bound morpheme, as described in the above sub-section,, are adjuncted to the right of the head sign. This structure pattern is retained in the current grammar implementation. During conversion, if an adjective phrase is present, np processing also involves a swap operation, the result of which is post-head positioning of adjective(s), while preserving input adjective phrase order of appearance. The output of swap operation is exhibited in example (Ex.13).

(Ex.13) KITCHEN LARGE BRIGHT
= *large, bright kitchen*

5. Future Research

The converter in its current implementation receives input from a shallow statistical parser for Greek which provides rough structural descriptions which do not carry extensive semantic information. The leaves of the so created structures contain feature descriptions which derive from a morphology based lexicon. In order to match input strings to adequate GSL structural representations, there has been used list matching according to semantic properties that are not directly visible in the source chunks, but, in this way, they are properly generated in the target structures. However, many GSL dependant issues remain untouched. The next research target involves searching for solutions as regards integration of classifier use in structure formation. An example is provided by the various GIVE formations that incorporate the classifier for the object. The natural signer incorporates the classifier indicating the semantic class of the object into the movement for GIVE formation, a procedure that creates a number of different entries in the lexicon, all recognized as various actions of giving.

sign: GIVE-MONEY / flat (perceived as-) 2D object
sign: GIVE-ROUND-OBJECT / 3D object
sign: GIVE-PENCIL / thin (perceived as-) 2D object
sign: GIVE-BOOK / flat (perceived as-) 3D object

6. Conclusion

The computational grammar exploited by the converter certainly covers a limited number of phenomena, and also reveals many of the issues still requiring an adequate handling in respect to their implementation, in order to achieve fully annotated strings as to information carried by natural signing utterances. However, its architecture allows extensibility with respect to further rule coding, at low computational cost.

The implemented subset of grammar rules are derived from an extensive formal grammar of GSL that captures the generative properties of the language. This grammar is the product of theoretical linguistic analysis of natural language data, and provides its first formal description, covering all levels of representation (phonology, syntax and semantics).

At the current stage, implementation has disclosed the potential of adequately coding signing linguistic information to an extent that allows recognition of the produced utterance as part of the language.

Acknowledgements

The authors wish to acknowledge assistance of all the GSL research group of ILSP and the IVMS lab. This work was partially funded by the national project grant SYNENNOESE (GSRT: eLearning 44).

7. References

Antzakas, K. & Woll, B. (2002). “Head movements and negation in Greek Sign Language”. In I. Wachsmuth, & T. Sowa, (eds.), *Gesture and sign language in human-computer interaction*. Lecture Notes in Computer Science. 2298, pp. 193-196. Springer.

Carpenter, B. (1992). *The Logic of Typed Feature Structures*. Cambridge University Press.

Efthimiou, E. & Zombolou, K. (1995).“Linear recognition or resolving the Modern Greek word order issue in NLP”, *Studies in Greek Linguistics*, pp. 635-644.

Efthimiou, E. (2006). Processing cumulative morphology information in GSL: the case of pronominal reference in a three-dimensional morphological system. In honour of G.Babiniotis, National and Kapodestrian University of Athens, in print.

Efthimiou, E., Vacalopoulou, A., Fotinea, S-E. & Steinhauer, G. (2004). “Multipurpose Design and Creation of GSL Dictionaries”. In Proc. of the Workshop on the Representation and Processing of Sign Languages “*From SignWriting to Image Processing. Information techniques and their implications for teaching, documentation and communication*”, Lisbon, Portugal, pp.51-58.

Fischer, S. D. (1996). The role of agreement and auxiliaries in sign language. *Lingua*. Special issue on "Sign linguistics: Phonetics, phonology and morpho-syntax" 98: 1-3, 103-119.

Fotinea, S-E., Efthimiou, E., & Kouremenos, D. (2005). “Generating linguistic content for Greek to GSL conversion". In Proc. of the HERCMA-2005 Conference, Sept. 22-24, 2005.

Hanke, T. (2002). HamNoSys in a sign language generation context. In Schulmeister, R., Reinitzer, H. (Eds) *Progress in sign language research*. In honor of Siegmund Prillwitz. *International Studies on Sign Language and Communication of the Deaf*; 40. Hamburg, 249-264.

Karpouzis, K., Caridakis, G., Fotinea, S-E. & Efthimiou, E. (2005). Educational Resources and Implementation of a Greek Sign Language Synthesis Architecture. *Computers and Education*, Elsevier, online since 09/05.

Marshall, I. & Safar, E. (2005). “Grammar Development for Sign Language Avatar-Based Synthesis”. In Proc. of the *3rd International Conference on Universal Access in Human-Computer Interaction* (UAHCI 2005), Las Vegas, Nevada, USA.

Prillwitz, S., Leven, R., Zienert, H., Hanke, T. and Henning, J. (1989). *HamNoSys. Version 2.0. Hamburg Notation System for Sign Language: An Introductory Guide*, Hamburg: Signum Verlag, 1989.

Sapountzaki, G. (2005). Markers of Tense, Aspect, Modality and Agreement as possible auxiliaries in Greek Sign Language. Doctoral Dissertation, University of Bristol.

Shieber, S.M. (1992). *Constraint-Based Grammar Formalisms*. MIT Press, Cambridge, Massachusetts.

Recognizing Hand Gestures Using a Fuzzy Rule-Based Method, and Representing them with HamNoSys *

Antônio C. R. Costa[1], Graçaliz P. Dimuro[1] and Benjamin C. Bedregal[2]

[1]Programa de Pós-Graduação em Informática, Universidade Católica de Pelotas, Pelotas, Brazil
[2]Departamento de Informática e Matemática Aplicada, Universidade Federal do Rio Grande do Norte, Natal, Brazil
{rocha,liz}@atlas.ucpel.tche.br, bedregal@dimap.ufrn.br

Abstract

This paper introduces a fuzzy rule-based method for the recognition of hand gestures acquired from a data glove, and a way to show the recognized hand gesture using the graphical symbols provided by the HamNoSys notation system. The method uses the set of angles of finger joints for the classification of hand configurations, and classifications of segments of hand gestures for recognizing gestures. The segmentation of gestures is based on the concept of "monotonic" gesture segment, i.e., sequences of hand configurations in which the variations of the angles of the finger joints have the same tendency (either non-increasing or non-decreasing), separated by reference hand configurations that mark the inflexion points in the sequence. Each gesture is characterized by its list of monotonic segments. The set of all lists of segments of a given set of gestures determine a set of finite automata that recognize such gestures. For each gesture, a sequence of HamNoSys symbols representing the reference hand configurations of the gesture is produced as an output.

1. Introduction

Fuzzy sets were introduced in 1965 by Zadeh (1965) for representing vagueness in everyday life, providing an approximate and effective means for describing the characteristics of a system that is too complex or ill-defined to be described by precise mathematical statements. In a fuzzy approach the relationship between elements and sets follows a transition from membership to non membership that is gradual rather than abrupt.

Fuzzy set theory is the oldest and most widely used theory for soft computing, which deals with the design of flexible information processing systems (Mitra and Pal, 2005). A fuzzy system implements a function (usually nonlinear) of *n* variables, given by a linguistic description of the relationship between those variables.

Figure 1 illustrates the architecture of standard fuzzy systems. The *fuzzificator* computes the membership degrees of the crisp input values to the linguistic terms (fuzzy sets) associated to each input linguistic variable. The *rule base* contains the inference rules that associate linguistic terms of input linguistic variables to linguistic terms of output linguistic values. The *information manager* is responsible for searching in the rule base which rules are applicable for the current input. The *inference machine* determines the membership degrees of the output values in the output sets, by the application of the rules selected in the rule base. The *defuzzificator* gives a single output value as a function of the output values and their membership degrees to the output sets.

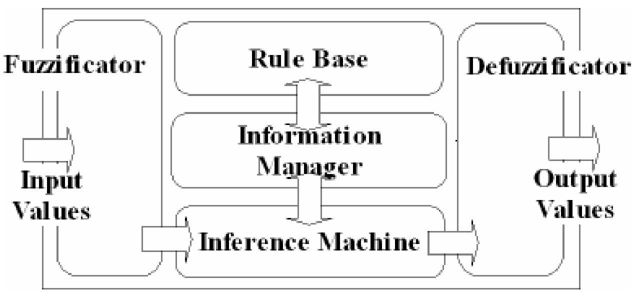


Fig. 1. Architecture of standard fuzzy systems.

The HamNoSys notation system (Prillwitz, S. et al.; 1989) is a graphical system for the symbolic representation of linguistic features of sign languages. In particular, it has a flexible set of symbols for the representation of hand configurations.

In this paper, we propose a *fuzzy rule-based method* for the recognition of hand gestures acquired from a data glove, and the representation of the recognized gestures as sequences in HamNoSys. We apply the method to the recognition of a sample hand gesture of LIBRAS, the *Brazilian Sign Language* (Brito, 1995).

The paper is organized as follows. In Sect.2, we introduce our fuzzy rule-based method for hand gesture recognition. A case study is discussed in Sect.3, with the recognition of a LIBRAS hand gesture. Section 4 shows the procedure for the automatic representation of hand gestures in HamNoSys. Section 5 is the Conclusion.

2. The Fuzzy Rule Based Method for Hand Gesture Recognition

The idea is to recognize some hand gestures with data obtained from a data glove with 15 sensors, as shown in Fig. 2. It is assumed that the data has been smoothed for jitter noise. The fingers are labelled as: F1 (little finger), F2 (ring finger), F3 (middle finger), F4 (index finger) and F5 (thumb). The joints in the fingers are labelled as J1 (the knuckle), J2 and J3, for each finger. Separations between fingers Fi and Fj are labelled as Sij.

Since any hand gesture can be represented as a sequence of frames, a hand gesture using a data glove is represented as a sequence of hand configurations, one for each discrete time instant. That is, at each time instant, the data glove sensors should provide the set of angles of joints and finger separation that characterizes a hand configuration.

In order to simulate this data transfer, a generator of hand configurations was implemented, generating monotonic sequences of handshapes: at each instant one hand configuration represented by a tuple of angles corresponding to each sensor shown in Fig. 2:

```
((F1J1,F1J2,F1J3), S12, (F2J1,F2J2,F2J3), S23,
 (F3J1,F3J2,F3J3), S34, (F4J1,F4J2,F4J3), S45,
 (F5J1,F5J2,F5J3) )
```

Given a hand configuration *c* and a sensor *s*, denote the value of a sensor angle by *s*(*c*), e.g., F1J1(*c*), S45(*c*) etc.

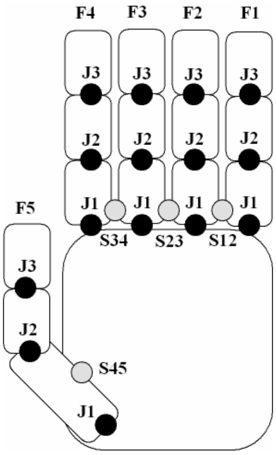


Fig. 2. Localization of sensors in the data glove.

2.1. Fuzzification

To each sensor corresponds a linguistic variable, whose values are linguistic terms representing typical angles of joints and separations. For the joints in the fingers (linguistic variables F1J1, F1J2, F1J3 etc.) the linguistic terms are: STRAIGHT (St), CURVED (Cv) and BENT (Bt). For the separations between fingers F1 and F2, F2 and F3, F4 and F5 (linguistic variable S12, S23, S45), the linguistic terms are: CLOSED (Cd), SEMI-OPEN (SOp) and OPEN (Op). For the separations between fingers F3 and F4 (linguistic variable S34), the linguistic terms are: CROSSED (Cr), CLOSED (Cd), SEMI-OPEN (Sop) and OPEN (Op). Figures 3, 4, 5, 6 and 7 show the fuzzification adopted for those variables.

2.2. The Recognition Process

The hand gesture recognition process is divided into four steps: (1) recognition of finger configurations, (2) recognition of hand configurations, (3) segmentation of the gesture in monotonic hand segments and (4) recognition of the sequence of monotonic hand segments.

For the step 1 (*recognition of finger configurations*), 27 possible finger configurations are considered. These configurations are codified in the following format: XYZ, where X is the value of the linguistic variable corresponding to the first joint J1, Y is the value of the linguistic variable corresponding to the second joint J2 and Z is the value of the linguistic variable corresponding to the third joint J3. For example, StStSt is used to indicate that the three joints are STRAIGHT, StCdCd indicates that the first joint is STRAIGHT whereas the others are CURVED etc.

The hand configuration is the main linguistic variable of the system, denoted by HC, whose linguistic terms are names of hand configurations, which names are application dependent. For instance, in Sect. 3, names of Brazilian Sign Language (LIBRAS) hand configurations (see Fig. 9) were used for such linguistic terms.

The 27 possible finger configurations determine 27 inference rules that calculate membership degree of each finger to each configuration. For example:

```
If F4J1 is STRAIGHT and F4J2 is CURVED and
   F4J3 is CURVED
Then F4 is StCdCd
```

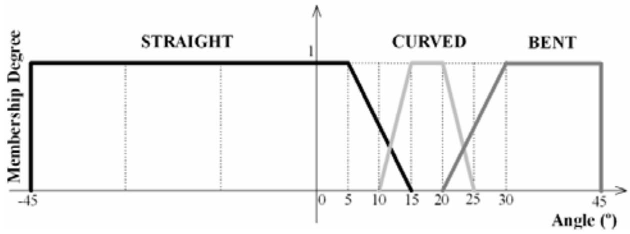


Fig. 3: Fuzzification of the linguistic variable of the joint F5J2 in the thumb finger F5.

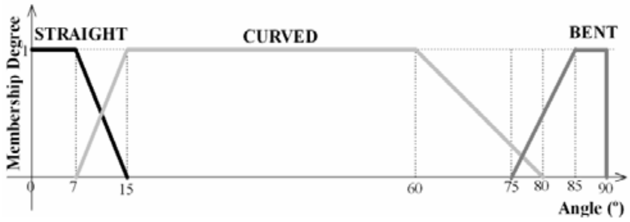


Fig. 4. Fuzzification of the linguistic variables of remaining finger joints

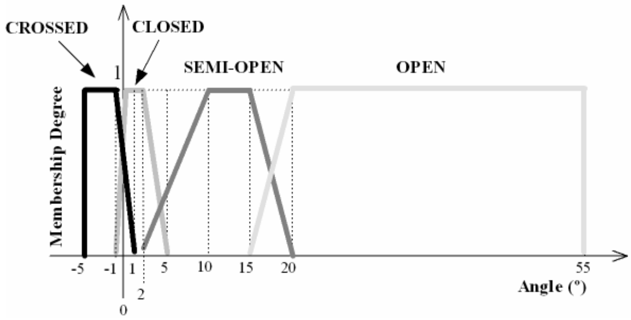


Fig. 5. Fuzzification of the linguistic variable of the separation S34 between the middle finger F3 and the index finger F4.

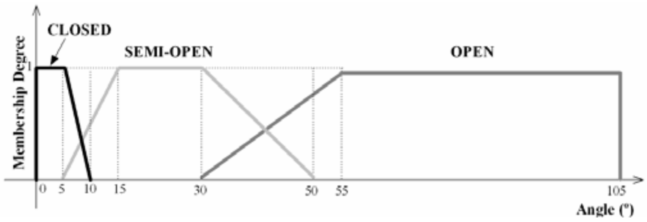


Fig. 6. Fuzzification of the linguistic variable of the separation S45 between the index finger F4 and the thumb finger F5.

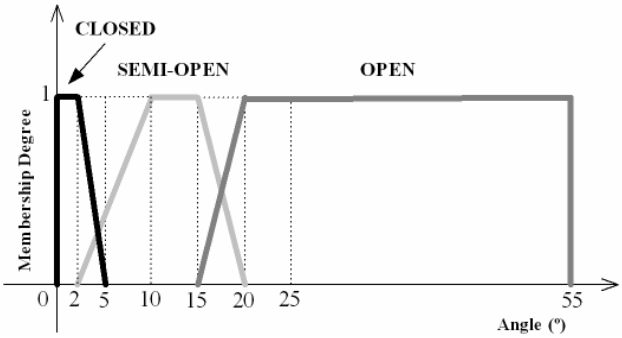


Fig. 7. Fuzzification of the linguistic variables of the separations between remaining fingers.

The next step is 2 (*recognition of hand configurations*), where the hand configuration is determined, considering each finger configuration and each separation between fingers. For example, the rule for the hand configuration [G] of LIBRAS (see Fig. 9) is described below:

```
If F1 is BtBtSt and S12 is Cd and F2 is BtBtSt
   and S23 is Cd and F3 is BtBtSt
```

```
    and S34 is Cd and F4 is StStSt
    and S45 is Cd and F5 is StStSt
Then HC is [G]
```

We note that since the hanshape is recognized as a unity, no co-articulation problem arises for its recognition.

In 3 (*segmentation of the gesture in monotonic hand segments*), we divide each gesture in a sequence of *k* limit hand configurations $l_1,\ldots,l_k$, where l_1 is the initial configuration and l_k is the terminal configuration. The limit configurations are such that, for each *c* between l_i and l_{i+1}, and for each sensor *s*, $s(c)-s(l_i)$ has the same sign of $s(l_{i+1})-s(l_i)$, for $i=1,\ldots,k-1$ (a difference of 0 is compatible with both negative and positive signs).

The limit hand configurations are the points that divide the gesture into monotonic segments, that is, segments in which each sensor produces angle variations with constant (or null) sign. For each monotonic segment l_il_{i+1}, l_i and l_{i+1} are its initial and terminal hand configurations, respectively.

The procedure for step 3 is the following. To find any monotonic segment l_il_{i+1}, the next *n* configurations sent by the data glove after l_i are discarded, until a configuration c_{n+1}, such that the signs of $s(c_{n+1})-s(c_n)$ and $s(c_n)-s(l_1)$ are not the same (or, c_{n+1} is the last configuration of the gesture). Then, c_n (resp., c_{n+1}) is the terminal hand configuration l_{i+1} of the considered monotonic segment, and also coincides with the initial configuration of the next segment $l_{i+1}l_{i+2}$ (if there is one). The process starts with $l_1 = l_1$, which is the initial gesture configuration, and is repeated until the end of the gesture, generating the list of *k* limit hand configurations.

In 4 (*recognition of the sequence of monotonic hand segments*), the recognition of each monotonic segment l_il_{i+1} is performed using a list of reference hand configurations $r_1,r_2,\ldots,r_m$ that characterizes the segment, where r_1 and r_m are the initial and terminal hand configurations of the segment, respectively. A monotonic segment is recognized by checking that it contains its list of reference hand configurations. The process is equivalent to a recognition based on a linear finite automaton, where $l_i = r_1$ and $l_{i+1} = r_m$ and the transition function is shown in Fig. 8.

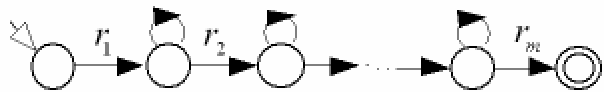


Fig. 8. Automaton for the recognition of monotonic segments.

3. Case Study: a Hand Gesture of LIBRAS

LIBRAS is the Brazilian Sign Language. As in any other sign language, the main parameters that characterize its phonological units are: the configurations of the hands used in the gestures, the main spatial location (relative to the persons who is signing) where the movements of the gestures are performed, the different movements (of the fingers in the hand, of the hands and arms in the space, of the whole body) that constitute the gesture, the facial expressions that express different syntactic, semantic and pragmatic marks during the production of the signs etc.

To support that recognition process, a reference set of hand configurations is usually adopted, driven either from the linguistic literature on sign languages, or dynamically developed by the experimenters with an ad hoc purpose. For our purposes, we have chosen a standard set of hand configurations (some of them shown in Fig. 9, taken from the linguistic literature on LIBRAS (Brito, 1995).

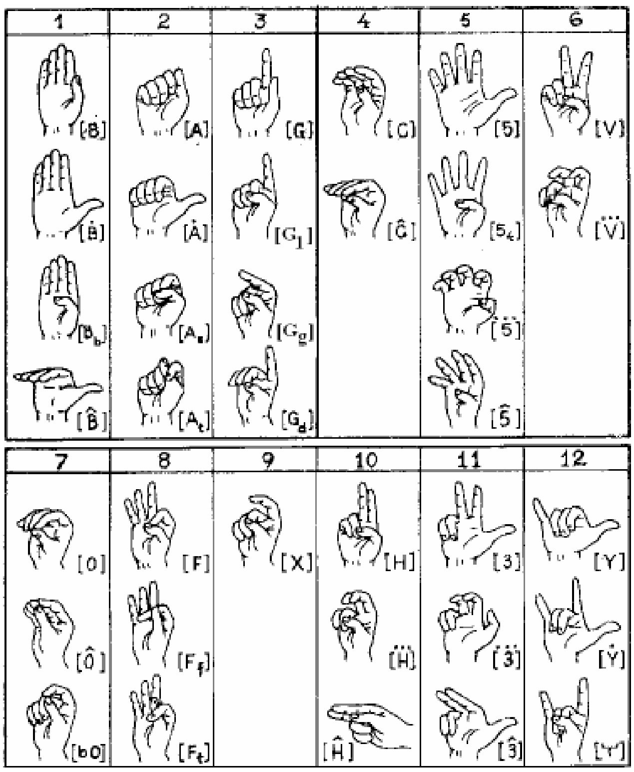


Fig.9. Some LIBRAS hand configurations.

Since we take the set of hand configurations from the literature, our method requires that each sign be thoroughly characterized in terms of its monotonic segments and the sequences of hand configurations that constitute such segments, and that the identification of the monotonic segments and hand configurations be manually provided to the system. Of course, a capture device such as a data glove can be used to help to identify the typical values of the angles of the finger joints, but the final decision about the form of the membership functions that characterize the linguistic terms used in the system has to be explicitly taken and manually transferred to the system.

We illustrate here the application of the method by the definition of the necessary parameters for the recognition of the hand gestures that constitute the sign CURIOUS, in LIBRAS. CURIOUS is a sign performed with a single hand placed right in front of the dominant eye of the signer, with the palm up and fingers oriented in the forward direction. The initial hand configuration is the one named [G1] in Fig. 9. The gesture consists of the monotonic movement necessary to perform the transition from [G1] to [X] and back to [G1] again, such movements been repeated a few times (usually two or three). Thus, a possible analysis of the hand gestures that constitute the sign CURIOUS in LIBRAS is:

Initial configuration: [G1]
Monotonic segment S1: [G1]-[G1X]-[X]
Monotonic segment S2: [X]-[G1X]-[G1]
State transition function for the recognition automaton: see Fig. 10.

To support the recognition of the monotonic segments of CURIOUS, we have chosen to use one single intermediate hand configuration, [G1X]. It is an intermediate configuration that does not belong to the reference set (Fig. 9) and whose characterization in terms of the set of membership functions for linguistic terms was defined in an ad hoc fashion, for the purpose of the recognition of CURIOUS. Together with [G1] and [X], it should be added to the list of hand configurations used by the recognition system.

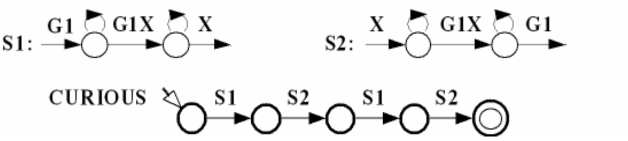


Fig. 10. Automaton for the recognition of hand gestures of the sign CURIOUS.

4. Representation of Hand Gestures in HamNoSys

HamNoSys (Hamburg Notation System) is a graphical system for the representation of linguistic features of sign languages. In particular, it has a flexible set of symbols for the representation of hand configurations.

Figure 11 presents the HamNoSys representation of some LIBRAS hand shapes.

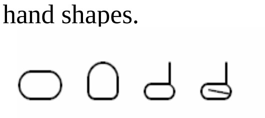


Fig. 11. The [A], [B], [G] and [G1] hand shapes respectively represented in HamNoSys.

It is easy to see how one can associate in a table the lists of features characterizing hand configurations with their representation in HamNoSys, so that the hand gesture recognition system can produce as output a sequence of HamNoSys notations.

Figure 12 shows the representation of the hand gesture of the sign CURIOUS.

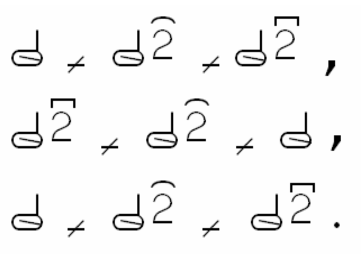


Fig. 12. The three monotonic segments of the LIBRAS hand gesture of the sign CURIOUS represented in HamNoSys.

Notice the following conventions in Fig. 12: the reference hand configurations within a monotonic gesture segment are separated by the symbol while the segments themselves are separated by the comma and ended by the point ; the initial configuration is [G1]; the other hand configurations are variations of [G1], denoted by indicating the finger whose configuration varies:

 indicates that the index figure is curved, while

 indicates that the index figure is bent.

Notice also that the order for the enumeration of fingers used in this paper (Fig. 2) is different from that adopted in HamNoSys.

5. Conclusion and Final Remarks

We presented a fuzzy rule-based for the recognition of hand gestures. The method is highly dependent on a detailed previous analysis of the features of the gestures to be recognized, and on the manual transfer of the results of that analysis to the recognition system. This makes it suitable for the application to the recognition of hand gestures of sign languages, because of the extensive analysis that linguists that have already done of those languages.

Prototypes of a random gesture generator and of the gesture recognizer were implemented in the programming language Python. In the fuzzification process, we considered only trapezoidal fuzzy sets and the minimum (or Gödel) t-norm, motivated by simplicity. The output of the recognition system was fed into a simple translator able to render the recognized hand gestures as they are annotated in the HamNoSys notation system.

Future work should develop in two directions: the recognition of arm gestures, by including the analysis of the angles of arm joints; the application of the recognizer to support computer systems controlled by pre-defined sets of signs, phonologically specified in HamNoSys.

6. References

Al-Jarrah, O., Halawani, A. (2001) Recognition of gestures in arabic sign language using neuro-fuzzy systems. *Artificial Intelligence* **133**: 117-138.

Bimber, O. (1999) Continuous {6DOF} gesture recognition: A fuzzy-logic approach. In: *Proc. of 7th. Intl. Conference in Central Europe on Computer Graphics, Visualization and Interactive Digital Media* (pp. 24-30}.v.1

Binh, N.D., Ejima, T. (2005) Hand gesture recognition using fuzzy neural network. *In: Proc. of ICGST Intl. Conf. Graphics, Vision and Image Proces.* (pp. 1-6), Cairo.

Brito, L.F. (1995) *Por uma Gramática de Línguas de Sinais.* Rio de Janeiro: Tempo Brasileiro.

Fang, G., Gao, W., Zhao, D. (2004) Large vocabulary sign language recognition based on fuzzy decision trees. I*EEE Trans. Syst., Man and Cybern*. **34**: 305-314

Hanmandlu, M., Yusof, M.H.M., Madasu, V.K. (2005) Off-line signature verification and forgery detection using fuzzy modeling. *Pattern Recognition* **38:**341-356.

Kwak, K., Pedrycz, W. (2005) Face recognition using a fuzzy fisherface classifier. *Pattern Recognition* **38**: 1717-1732.

Lin, C.T., Lee, C.S.G. (1996) Neural Fuzzy Systems: A neuro-fuzzy synergism to intelligent systems. Upper Saddle River: Prentice Hall.

Mitra, S., Pal, S.K. (2005) Fuzzy sets in pattern recognition and machine intelligence. *Fuzzy Sets and Systems* **156**: 381-386.

Prillwitz, S. et al. (1989) HamNoSys: Hamburg Notation System for Sign Languages. Hamburg: Signum.

Su, M. (2000) A fuzzy rule-based approach to spatio-temporal hand gesture recognition. *IEEE Transactions on Systems, Man and Cybernetics, Part C* **30**: 276-281

Zadeh, L.A. (1965) Fuzzy sets. *Information Control* **8**: 338-353.

[*] This work is partially supported by FAPERGS and CNPq (Proc. 470871/2004-0, Proc. 470556/2004-8).

Analysis of the different methods to encode SignWriting in Unicode

Guylhem Aznar, Patrice Dalle and Clément Ballabriga

Équipe TCI and Master Camsi

IRIT-UPS, 118 route de Narbonne, 31062 Toulouse Cedex 4
E-mail: aznar@irit.fr, dalle@irit.fr, ballabriga@7un.net

This paper lists, evaluates and discuss the solutions to encode SW in Unicode.

Abstract: SignWriting is the most complex and popular writing formalism for sign languages. Unicode is the most popular encoding of characters aimed at unifying the various language-oriented encodings into a single format supporting every human language. This paper focuses on the first functional layer, which gives a correspondence between a SignWriting sign and a series of bytes. This is one of the prerequisites to represent a sign language electronically. The different possibilities to encode a given SignWriting sign are evaluated and compared on different criteria : the Unicode space requirements, the number of bytes the storage will require, the mathematical complexity and the side advantages offered. Keeping as much as possible of the information on how signs are written and entered, and offering capabilities to easily compare the symbols that compose these signs is also considered, so that the encoding can serve to study and compare how SignWriting is written. A reference encoding is then proposed, to serve as a basis for the next layers. Other bi-dimensional writing formalisms, currently not supported by Unicode, are considered to extend the presented work.

1. SIGNWRITING

A sign language *sign*, corresponding to a meaning, is transcribed in a SignWriting (SW) *sign*, composed of symbols, positioned on a 2D canvas called a *signbox* (Sutton, 1995).

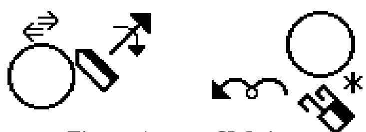


Figure 1 : two SW signs

Symbols correspond to static or dynamic positions or movements of the human body, and are described in the SymbolBank norm from the IMWA (Sutton, 2004). An analysis of SSS-2004 shows 25 973 symbols, divided into 8 categories, 10 groups, 50 elements, 5 variations, 6 fillings and 16 rotations.

Choosing a list of symbols from the SSS, and positioning them into a 2D signbox, results in a infinite number of combinations.

2. UNICODE

Unicode is simply an assignment of characters into code points. Unicode currently offers $2^{20}+2^{16}$=1 114 112 codes, split into 17 planes of 2^{16} = 65 536 codes. Only 100 000 characters have been assigned so far, i,e, 10% of the available code space. The first plane, also called Plane 0 , is used for existing encodings, to allow direct compatibility. It features a private area, a concept inherited from Asiatic encodings, used by systems or applications which must encode non standard characters. Plane 1 is used for ancient languages, mathematical and numerical symbols, and Plane 2 for rare, mostly historic, Chinese characters. Plane 14 currently contains non-recommended language tag characters and variation selection characters. Plane 15 and Plane 16 are fully reserved for private use. Unicode simply assigns a unique number to each character. But file storage, transfer and processing require handling these numbers following a mapping method. Unicode offers different ways to do such mappings, depending on constraints such as available storage space, compatibility requirements and interoperability, through various UTF and UCS. mappings. UTF-32 is the best choice when storage space or compatibility are less important than software uniformation, and will be used by default in this paper in an hexadecimal transliteration U+$X_1X_2X_3X_4$, where X_n is the n^{th} byte of an hexadecimal value X.

UNICODE ENGINES

Unicode does not deal with fonts : it simply matches first bits and codes, following a mapping like UTF-32, then codes and characters, following the standard arrangement of Planes. There is no bijection between the code and its graphical representation called “glyph”, unlike in traditional encodings such as ISO 8859-15 “Latin 9”: there are many ways to display a similar glyph. Matching one or more characters with a glyph is the job of the Unicode engine. For example, the French glyph “è” can be obtained through a single character called “LATIN SMALL LETTER E WITH GRAVE” which is given code U+00E8. Yet the same glyph can be displayed with the two characters “LATIN SMALL LETTER E” and “MODIFIER LETTER LOW GRAVE ACCENT”, respectively U+0065 and U+02CE. The latter could even be replaced by “COMBINING GRAVE ACCENT” U+0300 ! Such grammar, required to compose a glyph from characters, is called an Unicode engine (Fanton, 1998). There are many existing Unicode engines. The engine uses a font to represent the glyphs. There are currently less than ten “pan-Unicode” fonts, i.e. capable of supporting most of the glyphs Unicode can offer.
Some languages such as Arabic or Devanagari require a specific treatment of the glyphs. For example, in the

Arabic alphabet, most glyphs have four allographs, depending on the position of the letter in a word : isolated, initial, median, final. Such post-treatment of the glyphs into graphs is done by the Unicode engine.

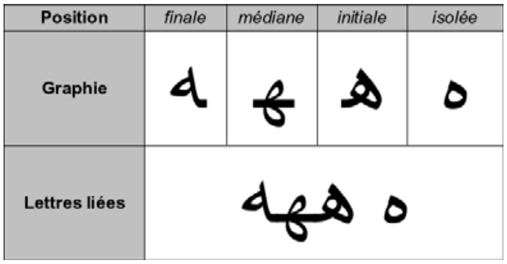


Figure 5 : the Arabic letter “hâ”, the words “hâ hâhâhâ”

3. USING UNICODE FOR SW ENCODING

SW is currently encoded in SWML (Da Rocha & coll, 2001). Unicode encoding can take advantage of the previously presented properties of Unicode : symbols encoding, and symbols positioning can be studied as separate problems, with the reconstruction left to the Unicode engine, which will require a grammar. In each approach, the following criteria must be considered : integration into the operating systems, minimization of the storage space required, minimization of the mathematical cost of the Unicode engine algorithm in CPU time, respect of the Unicode standard.

ENCODING THE SYMBOLS

The first problem is matching each symbol into a unique number. From that code, as explained before, various mappings will be available to encode the number into bytes. Only two solutions are possible, depending on the importance of the aforementioned criteria: a) minimization of the storage space : “sequential” approach, where symbols are not sorted into groups with special meanings, but simply follow a sequence without any given order b) minimization's of the CPU time : “bitwise” approach, where symbols are grouped. Each group corresponds to a given parameter of the symbol such as rotation, filling, etc. Each group corresponds to a bit field.

SEQUENTIAL APPROACH

Obviously, the easiest way to match a code to each of the 25 973 symbols is to proceed in sequence. This approach presents however a major problem : while SSS evolves on a yearly basis, inserting new symbols could logically only occur at the end of the block. Since symbols are organized following a given structure (categories, groups, elements, variations, fillings, rotation), this problem would be much more important than for other languages, especially for the software treatment of the given space: determining the parameter of a given code would require a case-by-case analysis for symbols outside the given space. And even in the given space, without the addition of any symbols, determining parameters would in the best case require modular arithmetic to perform range comparison and check whether a symbol belongs to a given group/category/etc.

BITWISE APPROACH

In order to minimize the CPU time, a bitwise approach could allow to easily find a symbol though a simple bit masking. However, this approach would increase the space used to encode the symbols. Let us start with an example tree of depth n, where for each level, each node can have the same number of leaf. In SW case, each level represents a parameter (c: category, g: group, e: element, v: variation, f: fillings, r: rotation). Following this simplification, each parameters can be represented as a set, function of the level i, called Ei. Encoding requires b bits, rounded to the next integer:

$$b_b = \sum_{i=0}^{n-1} (\ln_2(card(Ei)))\, bits$$

For example, the 6th level representing the rotation parameter corresponds to set E5 since we are starting at E0. Because there are 16 rotations, this set has 16 parts, and thus requires 4 bits. This bitwise approach thus has a mean costs of n-1 bits compared to the sequential approach. However, each of the 6 parameters c,g,e,v,f,r does not take an uniform amount of space : two different symbols can have different numbers of variations for example. Therefore, when encoding the variations into a fixed-length bit field, we must consider the worst case. Some space is wasted: it can be understood as identically sized boxes, which are as big as one of the box is filled, but globally are as empty as this case is rare.

COMPARING THE USE OF UNICODE SPACE

The cost of the sequential approach is known and fixed, and could easily fit in the Plane 2. The cost of the bitwise approach can be calculated with the previous formula. An analysis of SSS-2004 to calculate the card Ei for each of the 6 parameters reveals 23 bits are required. This means the bitwise encoding will require more than one Unicode per symbol since 2^{23} is 8 times greater than the total space offered by Unicode. The only possible solution is to use a sequence of 2 unicodes. It can come from a) an artificial extension of the Unicode space, which is contrary to the logic of the Unicode standard where each character must have its own code within the Unicode space or b) the use of modifier characters.

A MIXED BITWISE APPROACH WITH MODIFIERS

In the latter case, symbols are decomposed into a combination of parameters like for the “è” example : it will turn the chosen parameters into Unicode modifier characters. This would of course reduce the required Unicode space, but would let in exchange the storage space bear the equivalent cost of this simplification. At least two unicodes will be required anyway: if one parameter is made into a modifier, say variation for example, one Unicode is required for the $2^{23-\ln2(\text{card Ev})}=2^{17}$ codes corresponding to the remaining 5 parameters, and another Unicode is required for this modifier. Therefore, removing n parameters approximately results in a storage space requirements of 1+n unicodes. The approximation is due to the possible cases where “small” parameters could fit within a single code as cumulative modifiers, like ”rotated left with front face exposed”. The main interest of a mixed bitwise approach with modifiers is using a single unicode for the main part, to follow the Unicode logic of one code per character. For example, variation and rotation requiring respectively 6 and 4 bits could be made as a single modifier requiring 10 bits. The main part would then require a 13 bits unicode. Deciding which parameters will become modifiers will require a linguistic analysis of SW. Then, deciding whether modifiers will be encoded following a sequential or a bitwise approach will require another comparison, on speed, space and side benefits criterias.

COMPARING THE SPEED

A good example is calculating the speed to extract a parameter of the code of a given symbol such as the variation : text search operation, given the inter and intra personal variabilities, will frequently have to extract many parameters – and that for each symbol. Deciding to extract the "variation" parameter follows a simplification hypothesis, which will minimize the advantage of the best method, because there is always the same number of possibilities for the following parameters (rotation and filling). Extracting the variation of a given code when a sequential encoding is used could be done as :

```
int extract_variation_s(int code) {return ((c/nf*nr)%nv);}
```

Here *c* is the code, *nf* the maximal amount of fillings, *nr* the maximal amount of rotations, and *nv* the maximal amount of variations : *nf*=6, *nr*=16, *nv*=5. The most costly operations are 2 modular divisions, in 16 bits since we have less than 65 536 symbols. In the case of a bitwise encoding, the same function would be:

```
int extract_variation_b(int c){return ((c>>(br+bf))&7);}
```

where *br* is the amount of bits required to encode the rotation, *bf* the amount of bits required to encode the filling, and 7 is the decimal value of 111 binary, which is used to mask the 3 bits of variation. Here the most costly operations are a bit shifting on 32 bits and a bitwise "and" on 32 bits. A practical experimentation of an AMD Athlon XP 2400, 50 million operations take 1744 ms in the sequential approach, versus 292 ms in the bitwise approach. The bitwise approach with modifiers would represent an intermediate case where extraction of the parameters which are modifiers could require the use of modular operations if the modifiers are encoded in a sequential approach, while the other operations will be as fast as the full bitwise approach. These approaches should now be compared to SWML. The implementations may vary, but can be simplified to the minimal operation which will always be present when the variation will have to be extracted from a SWML-formatted symbol. This minimal operation is matching the pattern where the variation parameter is stored. The fastest possible way to perform that operation in C is with regular expression. Supposing the regexp is already compiled, to give a speed advantage to this approach:

```
regcomp(&preg,"<symbol[^>]*>[0-9]+-[0-9]+-[0-9]+-([0-9]+)-[0-9]+-[0-9]+</symbol>",REG_EXTENDED); regexec (&preg,SWML,2,tab,0);
```

This instruction is evaluated like the previous approaches on a AMD Athlon XP 2400. However, due to its low speed, it is only realized 50 000 times – it then takes 1788 ms. The Unicode sequential and bitwise approaches have been put in their worst possible configuration, and the SWML minimal step in its best possible configuration. The Unicode approaches still respectively perform 1025 times faster for the sequential approach, and 6123 times faster for the bitwise approach.

COMPARING THE STORAGE REQUIREMENTS

For the sequential approach, one Unicode will be necessary for each symbol. For the bitwise approach, two unicodes will at least be necessary for each symbol – regardless whether modifiers are used or not. SWML requires 18 characters per symbol. In conclusion, the proposed methods will use from 6 to 18 times less space.

COMPARING ADDITIONAL BENEFITS

From a video recognition perspective, a bitwise approach would also offer the additional advantage of fuzzy completion : in the case where the specific symbol is not fully recognized, setting the bits to identify which parameters were recognized (ex: rotation, element, etc.) would be a first step – other parameters could be prompted to the user, or guessed depending on the context (signs previously used, etc). From a linguistic perspective, a bitwise approach would also ease lexicographic treatment of sign languages : for a new unknown symbol, the recognized parameters would be filled in the fuzzy completion, while the missing parameters (ex: a new element) could be temporarily assigned a code in one of the private use areas, until a linguist can review it. From a standardization perspective, leaving some empty space to add future SSS symbols would cost no more than the space being wasted by a bitwise approach, following the assumption that "empty" groups and categories are the most likely to be completed in the future.

ENCODING THE SYMBOLS POSITION

SWML currently does not save any order in which the symbols are entered to create a sign, the symbols are simply positioned in a 128x128 area. Yet saving the order of symbols entry could be used in lexical analysis of SW. Unicode only features composition methods, ie grammatical ways to create glyphs from characters which are composed. However, Everson (2002) estimated that 8% of the remaining writing formalisms not yet supported by Unicode would require innovative rendering methods – such as 2D positioning for Mayan and Egyptian hieroglyphs. Therefore, we consider preserving the order of symbols, and offering and extensible 2D positioning.

A POSITION AND NO RELATION

The positioning problem can be subdivided into 2 problems : positioning the symbols on a 2D signbox, and describing the relation between the symbols. SWML currently does not describe the relation between the symbols, while their relation can have various meanings such as an ordered sequence of movements, temporal co-occurrence, contact between body segments, etc. This relation is simply described by adding additional symbols, which are also positioned on the canvas. This simplifies the problem, removing the "relation" feature, but also removes information which could be used later on. For example, contact between symbols is not defined. Should it be defined as a relation, this property (contact or the lack of) between symbols could be preserved even during magnification or minimization of the sign. Likewise, manipulation of the symbols linked in a spatial sequence could take advantage of that property to automatically reposition the other symbols when the symbol initiating the sequence has been moved. Such relations could also be used to simplify the Unicode engine grammar.

POLAR OR CARTESIAN COORDINATES

Following SWML approach and using a dedicated code to position a symbol in the 128x128 signbox would only require 16 384 codes, from a partially used Unicode plane or a personal use area in the worst case. Reserved planes, offering 2^{16} coordinates, can even be used for a finer positioning in 4 096x4 096 with two unicodes. If no precision is necessary, a single reserved plane can be used, offering $2^{16/2}$ ie 256x256 scale, with a single Unicode. The simplest solution is to decide on a center, and give coordinates from that center. This is the solution currently used by SWML. It could be made to keep the sequence of entered symbols. A variation of that method is using a dichotomies positioning, which can save space depending on the precision needed. Yet since at least one Unicode will be used with any method, it has no interest and artificially complicates this solution.

RELATIVE COORDINATES

This solution removes any signbox size limit, while also keeping the starting symbol and the order of the sequence as entered by the user. However, the algorithm is complex, since it needs a step-by-step reconstruction taking into account the preceding step to construct the sequence of symbols.

COORDINATES GIVEN BY A FUNCTION

A parameter of the coordinates could be given to a function, encoded along, which would return the other position. The algorithm would be as complex as the function required to position each symbol, which could be following the sequence under which they where entered. This encoding would be best used with image recognition, to track body trajectory movements. However this would be the most complex solution, since it would at least require a fitting function. A simplified version of this approach could be used for relation operators, which would then be considered as functions.

COMPARISON

The speed costs are too complex to be calculated. But obviously, every proposed solution could be used to position symbols on the signbox, to preserve the order of the symbols, etc. In any case, the minimal cost will be 1 Unicode. With so many similarities and very little advantages, it seems evident that the simplest method should be chosen depending on the needs. The polar coordinates were initially favored, before inter and intra-personal variations had been identified. Its interest now seems very limited. The relative positioning requires a step-by-step reconstruction, which brings unneeded complexity. No approach will provide significant advantages in the positioning method, except in very specific cases. Therefore, the Cartesian coordinates, already used by SWML must be recommended. The function based positioning method should be limited to a) image and video interpretation, to trace trajectories and b) relation operators, should they be implemented.

4. PROPOSED UNICODE ENCODING

SYMBOL ENCODING

We propose to support both the sequential encoding and the mixed bitwise encoding. The pure bitwise encoding is not proposed because is does not follow the logic of Unicode standard, and therefore may not be accepted by the Unicode consortium. The sequential encoding could be immediately used to offer backward compatibility with existing SWML systems while offering a 1000 fold speed increase for parameter extractions and a 18 fold space saving. The mixed bitwise approach will only be evaluable when turning parameters into modifiers will be agreed. Following the example where the variation and the rotation are turned into modifiers, it will require two unicodes, but fit within one plane since $2^{13}+2^{10}<2^{16}$

This approach will also provide a 6000 fold speed increase for the remaining parameter extractions, and offer fuzzy editing capabilities for video recognition software. Even if the current Unicode policy is against giving codes to pre combined characters, such advantages could help the request.

INTEGRATIVE POSITIONING APPROACH

A simple, non optimised, grammar, is proposed, with 3 elements: the symbol SYM, the operators OP, the parameters PAR, taken from a reserved plane to indicate the position. Since a sign is a set of positioned symbols, it is terminated by the TER special operator:

```
sign ->partialsign TER
partialsign -> partialsign element | element
element -> SYM | OP
 OP -> PAR| PAR PAR | CONTACT | SEQUENCE
SYM -> (existing symbols)
```

This basic grammar could be further optimised. Yet it provides a very simple way to position 2D Unicode symbols at some coordinates P by default, without any operator, from 256x256 to 4 096x4 096. It also adds two sample operators previously suggested :CONTACT, to indicate whether two symbols are touching, and SEQUENCE, to indicate a sequence of movements. They could be used as symbols or as operators. For example, in a mixed bitwise approach with filling and rotation as modifiers, a sample “deaf” symbol coud be:

```
HEAD12 10 20 FINGER FILL2ROT3 10 25 CONTACTSYMBOL 10 22 T
HEAD12 10 20 FINGER FILL2ROT3 10 25 CONTACT T
```

The first approach requires 11 codes, the second 9 codes. with a 4 096x4 096 signbox – this could be further optimised using a 256x256 signbox with a single PAR. Each approach would then take respectively 9 and 7 codes. The same sign in SWML requires 360 codes, with only a 128x128 signbox- between 40 and 50 times more. Additional operators could be added with the help of linguists, for SW or other languages needing specific spatial management - such as Mayan hieroglyphs.

5. CONCLUSION

Unicode will bring serious speed and size improvements. Moreover, the integrative positioning approach could be applicable to 8% of the languages requiring it. Giving a code to each symbol is not a complicated task for either approach – it can be fully automated, and use the private areas for quick prototyping until a dedicated area has been granted. But officially giving a code require describing the character (symbol) and its properties, in details, which will be a long and complex task. Transforming symbols into operators, thus expressing relations, will also be a challenge for linguists. But then Unicode will bring a real grammar to SW, and offer interesting relational information which will be usable in the user interface.

6. REFERENCES

Sutton V., Gleaves R. (1995), SignWriter - The world's first sign language processor. La Jolla, CA: Ed. Center for Sutton Movement Writing,

Sutton V. (2004), International Movement Writing Alphabet. La Jolla, CA: Ed. Center for Sutton Movement Writing

Da Rocha Costa A., Dimuro G. (2001) A SignWriting-Based Approach to Sign Language Processing. Gesture Workshap 2001, London

Fanton M. (1998) Finite State Automata and Arabic Writing. Proceedings of the Workshop on Computational Approaches to Semitic Languages (COLING-ACL'98), Montréal, PQ

Da Rocha A., Dimuro G., De Freitas J. (2004) A sign matching technique to support searches in sign language texts. Actes du Workshop RPSL, LREC 2004 (pp 32—34) Lisboa, Portugal

Everson M. (2002) Leaks in the Unicode Pipeline. 21st International Unicode conference, Dublin, Ireland

Sign Language Learning through Asynchronous Computer-Mediated Communication (CMC)

Maria Mertzani
Centre for Deaf Studies, University of Bristol
8 Woodland Rd, BS8 1TN, Bristol, U.K.
E-mail: M.Mertzani@bristol.ac.uk

Abstract

Current research shows that CMC provides an excellent vehicle for L2 learning since it affords both teachers and learners to communicate in an authentic learning environment where negotiation of meaning in the target language can take place in the same way as in face-to-face interaction. As bandwidth networks become more developed, it is feasible to transmit sign language communication using digitised video. In this paper, I present SignLab, a virtual sign laboratory at the Centre for Deaf Studies (CDS), in Bristol University, U.K., developed through the use of 'Panda' software. It is an asynchronous videoconferencing system developed for the learning of British Sign Language. In this paper, I discuss how SignLab changes the concept of traditional sign language teaching and learning in terms of course delivery, tutors' and students' online roles, course material and online communication and collaboration. At the end, I propose a framework based on constructivist and learner-centred principles that teachers may consider applying when teaching online.

1. INTRODUCTION

Traditionally, sign language teaching and learning is heavily depended on visual simulations (e.g., animation, text-books and video) for one and main reason: sign language is visual language and learners, in order to learn it, need to view its motion and its all inhibited non-manual characteristics. Many programmes use text-based material, CD/DVD multimedia and analog or/and digital video for the instruction of sign language. Text-based material (e.g., textbooks and dictionaries) is not enough for the studying of sign language as they consist of pictures and drawings, which cannot express the full emotion and its four-dimensional form that is the feature of sign language (Fourie, 2000; Sagawa & Teaceuchi, 2002). Learners need to be shown how to execute a sign and how certain modulations affect the meaning of signs. However, these modulations are not presented in published books and thus, learners are rarely able to convey grammar and/or semantics (Hoemann, 1978).

Videotapes, CD/DVD-ROMs and animated material solve this problem by incorporating video or/and animating images of sign language. Animated signing characters (e.g., signing avatars) of 2D or/and 3D designs can represent sign language but they require advanced skills in graphic design. The easiest solution to these problems is the integration of digital video into sign language classrooms (Cormier & Carss, 2004). Digitised video is now broadly used in videoconferencing systems, which enable second language (L2) learning to take place online.

In this paper, I introduce SignLab, an online virtual classroom which functions through the use of Panda software exclusively designed at the Centre for Deaf Studies, in University of Bristol, U.K. By presenting its facilitating features during the teaching and learning of British Sign Language (BSL), the discussion will focus on its impact on (a) course delivery, (b) students' and tutors' online roles, (c) course material (e.g., activities and assessment), and (d) online communication and collaboration. From a pedagogical perspective, this papers attempts to present a framework for the virtual learning of sign language, which tutors may consider applying when teaching from distance.

2. VIDEOCONFERENCING IN VIRTUAL LEARNING

The focus of this paper is on desktop videoconferencing technology. This term is used to describe desktop computers connected into the Internet and, fitted with a web cam and appropriate software, they allows users to communicate visually and in the target language much in the same way as in face-to-face communication (Martin, 2005; Smyth, 2005).

Videoconferencing is widely used in virtual learning and its potential has been recognized by recent research in different educational settings, at local and international level (Martin, 2005; Wang, 2004). It bridges the instructional gap which is created by physical distance between teachers and students since both can meet "face-to-face" visually, in real time or in an asynchronous mode of communication. It transforms the educational experience of people of all ages and it can be integrated into any curriculum and at all stages of education (Martin, 2005).

Actually, videoconferencing is in between face-to-face and text-based contact as far as verbal and non-verbal clues are integrated or not. Paralinguistic cues such as head nods and facial expressions improve understanding, increase confidence and reduce isolation between the users who are physically separated from one another (Wang, 2004). Regarding the video component, it is pointed out the importance of eye gaze awareness, the ability to monitor the direction of participant's gaze and thus, the focus of attention (Dustdar & Hofstede, 1999).

Research shows that video-based discourse reduces the difficulties of comprehending a L2 because the learner's potential for comprehension is increased if the visual information is included in the presentation (Gruba, 2004). This means that video allows learners to understand more than their linguistic knowledge permits and thus, motivates their learning. Through the display of props, actions and interaction learners "(1) narrow interpretations when they observe physical settings, (2) validate tentative hypotheses when they make sense of action, and (3) judge emotional states when they see interaction" (Gruba, 2004: 52). By using digitised video, users can stop the flow of information over poorly understood areas and concentrate on these in order to achieve better understanding.

The interactivity of online video makes learning more effective since the learner can ask questions directly to the teacher or to his/her colleagues. In addition the teacher can intervene and correct learner's language even with non-verbal information like gestures (Hada et al., 2002). The learning experience is extended by recording the videoconferences and using the videos as resources (Martin, 2005). Both teachers and learners can reuse the conversational videos for editing and reviewing past lessons. In this case, teachers can grasp learners' mistakes that they might missed during videoconferencing and students can memorize questions or feedback that occurred during online conversations (Hada et al., 2002). The video clips can then be posted on the Internet for the use of others (Martin, 2005).

Furthermore, videoconferencing allows online markup - based collaborative correction. By using the original digitised video, teachers and learners can intervene in video sequence and add their comments and corrections. Such correction system has been developed by Hada et al. (2002), which they have called the Video-based Communicative Language Learning System (Viclle).

However, video transmission is still a challenging area. When the bandwidth of the network is limited, data can be lost and this creates problems in the quality of the conversation. The motion of videoconferencing is 'jerky' and nuances in facial expression are lost. Because of perceptual latency problems it is often the case that participants do not understand the language of their interlocutors (Kinginger, 1998). Problems in video quality also disturb "the trajectories of the hands and arms which are essential in recognising sign language" (Ashourian et al., 2005: 1090). Therefore, the bandwidth required for real-time video transmission of sign languages needs to be greater that the bandwidth available on current networks (Saxe & Foulds, 2002). However, videoconferencing located within a Local Area Network (LAN) connection, such as within a university department, enables higher bandwidth that means higher-speed video transmission (Ryan et al., 2000; Smyth, 2005; Wang, 2004). Moreover, the small video windows on participants' computers screens do not allow full view of users' working environment. Information about who is sitting next to a user or who is not is limited since the camera is positioned on the front of participants' screens and the context cannot be fully viewed (Dustdar & Hofstede, 1999).

This restricted communicative space has resulted in the development of new linguistic and sociolinguistic sign language practices (Muir & Richardson, 2005). It is found that Deaf individuals modify their signing within this new space. For example, some signs usually produced above or at the waist level in the videoconferencing environment are produced with the hands almost at the chin level (Keating & Mirus, 2003).

3. INTRODUCING SignLab

Advances in desktop videoconferencing have enabled its use in sign language learning (Mertzani, 2005). Having a web cam, Internet connection and videoconferencing software, learners and teachers can meet "face-to-face" visually, and send or receive video information from their remote desktop computers. It enables Deaf and hearing to communicate and thus, to build an online sign language classroom, the SignLab. This classroom is based on a Local Area Network (LAN) connection within Centre for Deaf Studies, University of Bristol, U.K. where seven Apple Mac computers are networked and connected to a central 360 GB server.

Panda is the software installed in all computers and with which teachers and students work while being online. It allows very easy recording of video (and audio), which automatically compresses it into MPEG-4 format, a highly compressed format with minimal storage requirements and minimal time spent in waiting for compression and moving files between drives. By using Panda, students and teachers can film themselves signing, save the digitised video files in the server and share their work with each other or with other provisional users (Cormier & Carss, 2004).

In particular, SignLab comprises an asynchronous videoconferencing system since teachers and students are not online simultaneously, in real time but and there is a delay of hours or days between messages and their replies. As any other CMC system, it encompasses file sharing for information exchange through the use of specific software, which handles the capturing, restoring and representation of interaction through video. This is a process, which enables the personalisation of learning (Smyth, 2005) and facilitates peer editing and collaboration (Peterson, 1997; Warschauer, 1997).

Teachers and students when logging in SignLab, work in separate folders, the home directories. In these folders everyone identifies his/her workplace when entering into SignLab. Everyone's home directory is on the server and whatever is on his/her desktop, is on the server too. In particular, there are three folders on the server: Staff Homes, Staff Private and Teaching Resources. Staff Homes contains the home directories of all teaching staff and can be used by staff and students to send files to each other. Within each person's home is a folder called Public, and within Public is a folder called Drop Box. Anyone can put a file into the Drop Box but only the owner can view the contents of his/her Drop Box.

Staff Private is a shared folder for all staff without access to students and contains the teaching materials that tutors create. Staff can read and write to this folder but students cannot access it at all. Teaching Resources is also a shared folder for all staff and students. However, in this folder staff can both read (open and view) and 'write' (record and edit) while students can only read the files. From the Staff Private folder teachers drop their material to the Teacher Resources folder where students access and work with it while being online.

Panda is used for the delivery of BSL courses. By either filming themselves or by digitising old VHS tapes, teachers and students store these materials into the Teaching Resources folder where they have access and can retrieve them at their convenience. Panda-produced video conversations relate to specific tasks that teachers assign to their students and the completion of a task is the starting point for an on-going SignLab conversation.

In addition, the activities that are used in the SignLab classroom are interactive, in the sense that they actively engage students to the learning process. By using Panda, students are often asked to film themselves signing the content of a video clip they have watched (e.g., transliterate tasks) or do a self-analysis of their own produced video clips. These activities are produced in BSL, thus creating an authentic sign language environment, where students are immersed into the target language.

While being online, tutors' and students' roles are changing. From SignLab experience it is shown that students are more independent and empowered for their own learning. They are recipients and have the control over the learning process. By watching over and over the video material, students correct themselves, imitate Deaf signers' signing and form forms of the target language. In this way, they are able to analyse their signing and realise their strengths and weaknesses (Mertzani, 2005). Each Panda window has basic video control buttons (play, stop, pause and rewind), so that students are able to watch the video by jumping to specific scenes (backwards and forwards), which are interesting or difficult to understand.

The main tutors' role is to answer questions and give feedback concerning unknown vocabulary, syntax and grammar. They usually check students' video signing and post their feedback to students' folders. Panda offers the possibility for an online mark-up assessment, similar to Hada et al. (2005) system introduced above. By using Panda, tutors can open students' original video clip, intervene in its sequence and add their comments and corrections. In this way, they grasp students' mistakes and assess better their BSL skills. On the other hand, students can memorize questions or feedback that occurred during online conversations.

This online communication and collaboration is one-to-one only. These conversations take place between teachers and students rather than between students themselves, but they can be teacher and/or student initiated. Therefore, one teacher or student is able to send his/her video message to another teacher or student only.

Many researchers have argued that CMC provides an excellent vehicle for L2 learning, based on the key premise that CMC affords teachers and learners to negotiate meaning while focusing on the linguistic part of language (Meskill & Antony, 2005). Some of the reasons cited for this assertion are: (a) increased reflection time; (b) more democratic participation; and (c) increased L2 production and discourse quality. From current practices and from preliminary research data (Mertzani, 2005), these reasons appear to apply to SignLab too.

Increased reflection time means that both teachers and students are afforded the needed time to attend to and process the target language, since CMC consists of 'written speech' where language forms are "visually immediate". For learners, it means that they have the opportunity to reflect upon and to look at the form and content of the online message as many times and for as long as they wish (Meskill & Anthony, 2005; Smith, 2003). For teachers, it means that they can detect learners' language, edit their responses and respond to the 'teachable' moments that rendered by the online conversation – see above the online mark-up assessment by using Panda - moments that in classroom time may not have been perceivable (Meskill & Anthony, 2005: 92). For this reason it is claimed that asynchronous is more beneficial than the synchronous CMC (Lamy & Goodfellow, 1999).

In this vein, research in spoken languages has indicated that extra time conversing online in the target language improves students' communicative competence, reading and writing skills (Sanchez, 1996). There is no yet similar research for sign languages, however it is possible to claim that more time on tasks may lead to sign language skills improvement. Furthermore, because of the scarcity of opportunities to hearing students to use sign language outside their classes in meaningful communication, SignLab is a useful tool for language access. Additionally, it is a comfortable environment for students, as they can watch their material and join in conversation whenever they feel ready.

Moreover, CMC is a less stressful environment for L2 learners especially for those who are traditionally silent or apprehensive producing verbal output in class. It is found that they increase their participation in online discussions compared to face-to-face in the L2 classroom (Jepson, 2005; Warscaheur, 1997). This is because CMC "(a) reduces social context clues related to race, gender, handicap, accent, and status … (b) reduces nonverbal cues, such as frowning and hesitating, which can intimidate people, especially those with less power and authority … [and] (c) allows individuals to contribute at their own time and pace"(Warschaeur, 1997: 473). Consequently, CMC enables learners from varying levels of L2 proficiency and ability to willingly experiment with forms of the target language and to assist one another during online activities (Jepson, 2005).

SignLab is proved to be a relaxed environment for students' learning. Although students are having a visual online communication with their tutors, they are less apprehensive, since they work only with their tutors and they are not exposing themselves to their colleagues. Thus, students feel more comfortable and relaxed, they experience less embarrassment by their mistakes and they are willing to produce more output than in their regular classes. Additionally, tutors' comments are addressed to just one student, so that only the student being corrected can watch the message.

Research has reported that learners develop more complex lexically and syntactically language in their online discussions, which covers a wide range of discourse functions similar to characteristics of oral and written language (Smith, 2003; Warschaeur, 1997). As I have already mentioned, online video communication has resulted in the change of people's signing. Such changes are observed during SignLab conversations. For example, students and teachers, before sending any message, orient themselves in front of the camera and adjust their signing in the visual field of it. They reorganize their sign space and modify sign location and orientation within this new space as well as repeat and slow down their signs. Some students produce video clips in order to check themselves signing. If the signing is not satisfactory (e.g., they are making mistakes while signing), they delete the video and try to produce a new one, avoiding making old signing errors.

4. A FRAMEWORK FOR ONLINE SIGN LANGUAGE PEDAGOGY

As any other CMC environment, SignLab changes teaching by focusing teachers' perspectives on a learner-centred design of instruction (Salaberry, 2000). There are important differences between a traditional and a CMC sign language classroom, as they resulted from current SignLab practices:

1. Recording and exchanging signing involves one person at a time (one-to-one communication).
2. The learning is student-centered rather than teacher-centered.
3. Students function in both initiating and responding roles (asking, giving information and negotiating meaning).
4. The learning is self-pacing and it can occur at any time and any place.
5. The teacher is the facilitator of students' learning rather than the content specialist.
6. The teaching is a constructive process rather than an instructive process.
7. Students work individually with different assignments as well as assessed individually.

SignLab experience shows the importance of a learner-oriented approach in order to match students' needs in their daily work (McAvinia & Hughes, 2003; Palloff & Pratt, 2003). Therefore, there needs to be a general agreement over new sign language pedagogy in terms of language learning methodologies that follow constructivist principles that are currently applied in L2 virtual learning environments. The adaptation of such approaches to videoconferencing, such as SignLab, "require[s] thoughtfulness, reflection and planning so it is probably wise to consider the use of a planning framework … for deciding which types of interactions might appropriately be [employed]" (Smyth, 2005:809).

Constructivistic models of learning call for specification and use of authentic and complex activities during the learning process so that students can perform the tasks by critically reflecting on them (Henze & Nejdl, 1998: 64). CMC should be used not so much to teach curriculum objectives in a different way, but rather to help students understand how their knowledge can be constructed by online collaboration practices (Kern et al., 2004). Consequently, there is the need for sign language tutors to agree upon an overall teaching and learning strategy, which can be adopted by all staff "and not left to the efforts of one or two academics and therefore seen as peripheral" (Gillepsie & McKee, 1999: 452).

The challenge for teachers is to integrate asynchronous CMC into sign language teaching. To present, there is no syllabus specially developed for SignLab and as a result, they follow the one they use at their regular BSL classes. Thus, online sessions and the material used in the SignLab should be carefully sequenced within a curriculum that follows the principles of constructivist methodologies.

This means the implementation of compulsory structure activities that promote online interaction (student-tutor, student-student). Research on L2 learning through CMC shows that jigsaw and decision-making tasks affect students' language acquisition, especially when structure activities are managed by the tutor; these are more likely to result to L2 learning (Smith, 2003; Paran et al., 2004).

Such tasks must be considered integrating into SignLab, although more research is needed to investigate their influences on sign language learning. In addition, the pedagogical design of these tasks must be based on the defining features of CMC environments (Salaberry, 2000; Skehan, 2003), in our case, SignLab environment.

The framework does not imply that tutors should reduce their contribution to managing online activities. On the contrary, tutors are able to work with students individually and develop a personal relationship with them, thereby understanding their needs and control their learning process. This corresponds to conclusions by other researchers (Stepp-Greany, 2002) that tutors play a significant role in a CMC technology instruction. Furthermore, changes in students' role need to be considered by tutors for reflecting on their teaching practices in order to facilitate students' sign language learning (Lam & Lawrence, 2002).

Additionally, the development of CMC environment should be based on the following seven hypotheses (Carol, 1998: 23-25):

1. The linguistic characteristics of L2 input need to be made salient.
2. Learners should receive help in comprehending semantic and syntactic aspects of linguistic input.
3. Learners need to have opportunities to produce L2 output.
4. Learners need to notice errors in their own output.
5. Learners need to correct their linguistic output.
6. Learners need to engage in L2 interaction for the negotiation of meaning.
7. Learners should engage in L2 tasks for maximizing their interaction.

Therefore, online teaching must consider adopting two types of tasks: (a) the "knowledge construction tasks" and (b) the "collaboration tasks". The first category involves tasks that promote learners' construction of sign language skills (receptive, expressive or both). Through these tasks students develop their knowledge by observing and modelling the language. In addition, these tasks comprise the starting point for "collaborative tasks" which constitute on-going discussions about the outcome of a "knowledge construction task". Through these tasks learners develop language by reflecting on the video recorded 'talk'. Students can ask questions, teachers can provide information and feedback (immediate or delayed), teachers and students can come to an agreement upon certain error types and learners can reflect on the feedback and on their own performance. Both tasks can be teacher and/or student initiated.

5. CONCLUSION

SignLab as described in this paper is the first virtual classroom for sign language learning. The last few years we have seen a dramatic expansion of Internet sites concerning online sign language learning and this trend will continue to occur as bandwidth for video transmission is developing. The emergence of such environments is challenging sign language teachers to consider their online roles and teaching strategies. We are still in the early stages and there remains the need for extensive future research. There are still many potential problems associated with the utilisation of SignLab, but more research will shed light into the online educational process.

However, SignLab applications do indeed create the necessity to develop an online pedagogy, including teaching and learning processes that are different from those occurring in traditional sign language environments, of which educators need to be aware. SignLab is a promising online learning tool, yet there is the need to learn more about it in order to unlock its potential for sign language learning.

6. BIBLIOGRAPHICAL REFERENCES

Ashourian, M., Enteshari, R., & Lambadaris, I. (2005). A Low Cost and Efficient Sign Language Video Transmission System. In R. Khosla, R. J. Howlett, & L.C. Jain (Eds.), *Knowledge-Based Intelligent Information and Engineering Systems: 9th International Conference, KES 2005, Melbourne, Australia, September 14-16, 2005, Proceedings, Part II*. Springer Berlin Heidelberg. http://www.springerlink.com/media/2g271r8tqgcwrn42wmf0/contributions/2/e/t/f/2etflbfqujv5rm8u.pdf

Carol, A.C. (1998). Multimedia CALL: Lessons to be Learned from Research on Instructed SLA. *Language Learning and Technology*, 2(1), 22--34.

Cormier, K., & Carss, P. (2004). *SignLab: How to Use Equipment and Software (Staff + Students)*. Bristol: Centre for Deaf Studies, University of Bristol.

Dustdsar,S., & Hofstede, G.J. (1999). Videoconferencing Across Cultures – A Conceptual Framework for Floor Control Issues. *Journal of Information Technology*, 14, 161--169.

Fourie, R.J. (2000). Efficiency of a Hearing Person Learning Sign Language Vocabulary from Media Versus Teacher. *Deafness and Education International*, 2(1), 45--60.

Gillepsie, J., & McKee, J. (1999). Does it Fit and Does it Make Any Difference? Integrating CALL into the Curriculum. *Computer Assisted Language Learning*, 12(5), 441--455.

Gruba, P. (2004). Understanding Digitised Second Language Videotext. *Computer Assisted Language Learning*, 17(1), 51--82.

Hada, Y., Ogata, H., & Yano, Y. (2002). Video-Based Language Learning Environment Using an Online Video-Editing System. *Computer Assisted Language Learning*, 15(4), 387--408.

Henze, N., & Nejdl, W. (1998). A Web-based Learning Environment: Applying Constructivist Teaching

Concepts in Virtual Learning Environments. In F. Verdejo & G. Davies (Eds.), *The Virtual Campus. Trends for Higher Education and Training* (pp. 61-77). London: Chapman and Hall.

Hoemann, H.W. (1978). *Communicating with Deaf People: A Resource Manual for Teachers and Students of American Sign Language.* Baltimore: University Park Press.

Jepson, K. (2005). Conversations -- And Negotiated Interaction -- In Text and Voice Chat Rooms. *Language Learning and Technology*, 9(3), 79--98.

Keating, E., & Mirus, G. (2003). American Sign Language in Virtual Space: Interactions between Deaf Users of Computer-Mediated Video Communication and the Impact of Technology on Language Practices. *Language in Society*, 32: 693--714.

Kern, R., Ware, P., & Warschauer, M. (2004). Crossing Frontiers: New Directions in Online Pedagogy and Research. *Annual Review of Applied Linguistics*, 24, 243--260.

Kinginger, C. (1998). Videoconferencing as Access to Spoken French. *The Modern Language Journal*, 82 (4), 502--513.

Lam, Y. & Lawrence, G. (2002). Teacher-Student Role Redefinition During a Computer-Based Second Language Project: Are Computers Catalysts for Empowering Change? *Computer Assisted Language Learning*, 15, 295--315.

Lamy, M.N. & Goodfellow, R. (1999). "Reflective Conversation" in the Virtual Language Classroom. *Language Learning and Technology*, 2, 43--61.

Martin, M. (2005). Seeing is Believing: The Role of Videoconferencing in Distance Learning, *British Journal of Educational Technology,* 36(3), 397--405.

McAvinia, C., & Hughes, J. (2003). Sharing Systems, Sharing Language: Designing and Working in a Virtual Department. *Computer Assisted Language Learning*, 16(5), 445--467.

Mertzani,M.(2005). Computer-Mediated Communication Technology in Sign Language Teaching and Learning. Its Impact on Students' and Teachers' Roles in the Classroom. Paper presented at the 38th Annual Meeting of BAAL, University of Bristol, September 15 - 17.

Meskill, C., & Anthony, N. (2005). Foreign Language Learning with CMC: Forms of Online Instructional Discourse in a Hybrid Russian Class. *System*, 33, 89--105.

Muir, L.J., & Richardson, I.E. (2005). Perception of Sign Language and Its Application to Visual Communications for Deaf People. *Journal of Deaf Studies and Deaf Education*, 10 (4), 390--401.

Palloff, R.M., & Pratt, K. (2003). *The Virtual Student. A Profile and Guide to Working with Online Learners.* San Francisco: Jossey-Bass.

Paran, A., Furneaux, C., & Summer, N. (2004). Computer-Mediated Communication in Distance MA Programmes: the Students' Perspective. *System*,32, 337--355.

Peterson, M. (1997). Language Teaching and Networking. *System*, 25(1), 29--37.

Ryan, S., Scott, B., Freeman, H., & Patel, D. (2000). *The Virtual University. The Internet and Resource-Based Learning.* London-Sterling (USA): Kogan Page.

Sagawa, H., & Takeuchi, M. (2002). A Teaching System of Japanese Sign Language Using Sign Language Recognition and Generation. In *International Multimedia Conference 2002. Proceedings of the 10th ACM International Conference on Multimedia, December 1-6* (pp 137--145). New York: ACM Press. http://doi.acm.org/10.1145/641007.641035

Salaberry, M.R. (2000). Pedagogical Design of Computer Mediated Communication Tasks: Learning Objectives and Technological Capabilities. *The Modern Language Journal*, 84, 28--37.

Sanchez, B. (1996). MOOving to a New Frontier in Language Teaching. In M. Warschauer (ed.), *Telecollaboration in Foreign Language Learning. Proceedings of the Hawaii Symposium. Honolulu* (pp. 145--163). Hawaii: Second Language Teaching and Curriculum Center, University of Hawaii.

Saxe, D.V., & Foulds, R.A. (2002). Robust Region of Interest Coding for Improved Sign Language Telecommunications. *IEEE Transactions on Information Technology in Biomedicine*, 6(4), 310--316.

Skehan, P. (2003). Focus on Form, Tasks, and Technology. *Computer Assisted Language Learning*, 16 (5), 391--411.

Smith, B. (2003). The Use of Communication Strategies in Computer-Mediated Communication. *System*, 31, 29--53.

Smyth, R. (2005). Broadband Videoconferencing as a Tool for Learner-centred Distance Learning in Higher Education. *British Journal of Educational Technology*, 36 (5), 805--820.

Stepp-Greany, J. (2002). Students Perceptions on Language Learning in a Technological Environment: Implications for the New Millennium. *Language Learning and Technology*, 16, 165--180.

Wang, Y. (2004). Supporting Synchronous Distance Language Learning with Desktop Videoconferencing. *Language Learning and Technology*, 8(3), 90--121.

Warschauer,M. (1997).Computer-Mediated Collaborative Learning: Theory and Practice. *The Modern Journal*, 81, 470--481.

Towards a Corpus-based Approach to Sign Language Dictionaries

Thomas Hanke

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

Abstract

This paper discusses those aspects of iLex, a sign language transcription tool, that are relevant to lexical work and the production of e-learning materials. iLex is built upon a relational database, and uses this strength to support the user in type-token matching by giving immediate access to all other tokens already related to a certain type. iLex features a number of classification schemes, both built-in and data-driven, to allow for the incremental process of identifying and describing the lexicon of a sign language. Data cannot only be exported to other transcription tools, but also into authoring systems for teaching materials. Finally, we speculate about the applicability of Zipf's Law for sign language corpora extrapolating from the current contents of the iLex database.

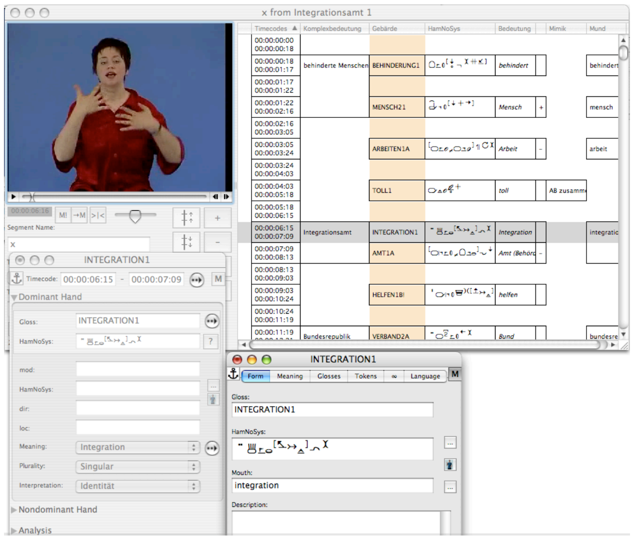
1. Introduction

Over the past fifteen years, our institute has produced a number of special-terminology dictionaries for German sign language. Having started with introspective data from small focus groups, we quickly moved to an empirical approach where informants were invited to report on their professional experience and to answer to a variety of elicitation settings. The signers' productions were recorded on video and later transcribed. These data provided the almost exclusive source for the dictionaries. A common starting point for all these dictionary projects was to use a list of concepts to be covered in each respective dictionary, usually defined from an educational point of view. The amount of video data collected and thereby the transcription effort needed were mainly determined by the size of this list and the number of informants available.

For these projects, we developed methods and tools to support the transcription and further analysis processes, especially type-token matching – a task much harder than for most spoken languages (Hanke et al., 2001). Moving away from written-language phrases or pictures as elicitation prompts towards semi-structured interviews and discussions, the transcription tools needed to become flexible enough to transcribe any signed discourse, not just short mainly sequential phrases. This now allows us to use the same tool named iLex both for lexicographic work and discourse analysis (Hanke, 2002b, for tools in our earlier work in discourse transcription with syncWRITER cf. Hanke, 2001).

2. Type-Token Matching

As sign languages have no written form, language resources for sign language often use “phonetic” notations, such as HamNoSys (Prillwitz et al., 1989 and Schmaling/Hanke, 2001, Hanke 2004). However, the current state-of-the-art for sign language notation is far away from being a full compensation for an orthography (Miller, 2001), which in general is the main access key to language data for written language as well as annotated speech. We therefore consider it essential for sign language corpus annotation to explicitly link tokens to lexical entities. The distinctive feature of our transcription tool is that it is built on top of a relational database

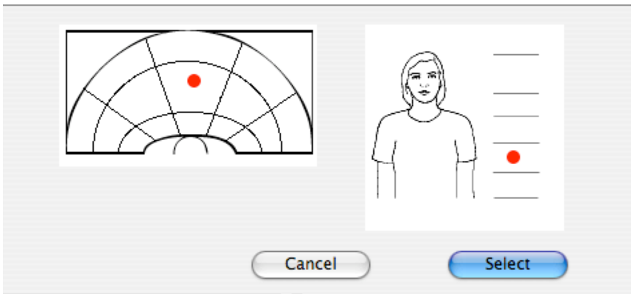


modelling tokens and types as different entities related to each other. I.e. stretches of signed discourse cannot only be tagged with text, e.g. glosses, but also as tokens related to one specific type.

The major advantage of this approach is that in the course of type-token matching, one can always review the video clips showing other tokens related to a candidate type. In addition, the relational model allows a multitude of search approaches to identify candidate types, e.g. by meaning, by gloss[1], by form, or by grammatical class.

Once the type for a token is identified, deviation in form from the type needs to be considered. In the case of grammatical modifications, such as inflection, the system suggests possible categorisations of the modification based on the assumed grammatical class of the type.

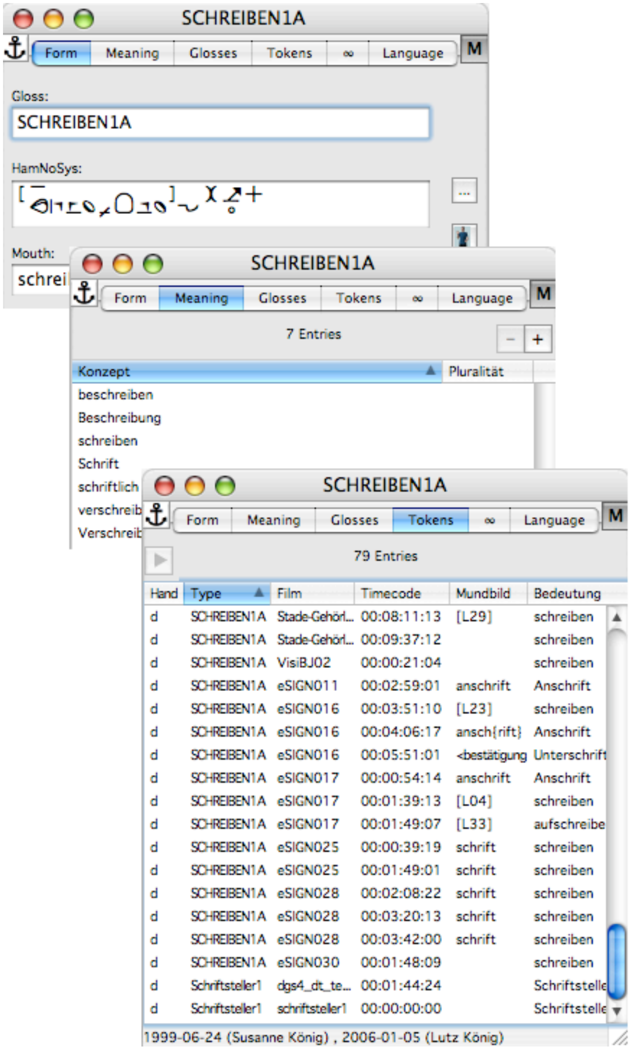
[1] We share the view of many researchers that glosses are convenient labels for types. It is of course always necessary to keep in mind the danger of using spoken language words for sign language types (cf. Pizzuto/Pietandrea, 2001), and even native signer team members report about various occasions where spoken language labels mislead them. The database approach however implies that token data are constantly reviewed from a number of perspectives, and in many cases glosses play no role so that it is our hope that such cases will be identified even in projects where, for budget reasons, not all transcription work can be independently reviewed.



Should a revision of the grammatical classification of a type render modification classifications of some related tokens invalid, the tool suggests these tokens to be reviewed.[2]

3. Type Hierarchy

iLex allows the transcriber to arrange types in a tree hierarchy. We currently use a four-level schema: On the top level, we describe the abstract images or ideas



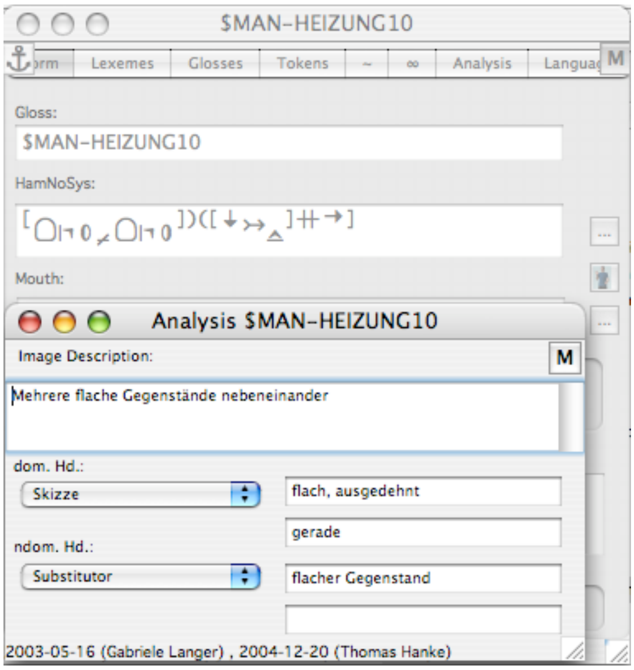
underlying many signs. The manual realisations of these images, i.e. certain forms, are found on the level below. This level is what most researchers would consider the sign inventory. Proper homonyms (non-polysemic) exist on this level: They share their surface form, but are derived from different images.

Forms can be assigned certain meanings, and there are numerous examples of signs with many different meanings. Conventionalised form-from-image/meaning pairs are notated on the third level, allowing e.g. their own glosses without obscuring their relation to other meanings that share form and image. In our model, this level corresponds to the lexicalisations level in a dictionary. For DGS, we can often notate a default mouth picture on this level.

On the lowest level, both forms and form/meaning pairs can be split up according to project-specific needs. E.g. projects using the transcription tools in order to produce e-learning materials currently use this level to assign alternative glosses they consider more appropriate in a certain didactic context. It is also possible to consider some modifications of types as separate entities and to use this level to introduce the dependencies, e.g. to introduce separate I and YOU if on the higher level only one (person) reference entity exists.

First experiments indicate that this hierarchical approach also has the potential to model the overlap of the type inventories of different sign languages. For this purpose, types can be attributed with languages that they appear attested for. Filters then allow the user to concentrate on the type inventory and associated tokens for one specific language only or to view data from a multitude of languages at the same time.

iLex allows the user to define irreflexive relations between types of a certain level. We currently use this to further analyse homonymic as well as close-neighbour form relations. In addition, types can be analysed with respect to the image production techniques used (Konrad et al. 2004).



[2] As the tool is used by different projects within our institute and as conceptualisations also change over time within one research group, grammatical classes and modifications they allow are modelled by the database as well, so that the grammatical model applied is not determined by the system, but the data. For special graphical editors needed to make data input more efficient (such as the one shown in fig. 2), a plug-in model has been implemented.

4. Representation of Form

Obviously, since the advent of digital video and especially in a database context such as iLex, where videoclips associated with tokens are immediately available, phonetic transcriptions no longer have an exclusive role to describe the form. But as long as is not possible to automatically search a video for certain features of a sign, phonetic data allows access to the data not otherwise possible. For tokens, HamNoSys as we use it in Hamburg certainly is an adequate form description. HamNoSys-compatible avatar software (Elliott et al. 2004) allows the transcriber to immediately verify HamNoSys notations written in iLex. Types, on the other hand, may require a different system.

While we use HamNoSys here as well with recent additions to allow underspecification or ranges of permissible handshapes, for example (Hanke 2002a), an easier way to abstract away from individual variations is highly desirable. Phonological models, however, still await the availability of large lexical databases in order to be verified.

5. Data Exchange

For a larger lab, where not only researchers work on transcriptions, but also students, a central database has the major advantage that transcription cannot go lost as people leave or students finish their exams. This cannot, however, mean that sharing data with the research community should become more cumbersome. ilex therefore provides export modules to transfer selected transcriptions to formats used by other researchers, such as ELAN (Crasborn et al. 2004) or SignStream (Neidle 2001). For read-only purposes, transcriptions can also be exported as QuickTime movies with subtitles or as scores in HTML format to be viewed with any browser. In cases where the original video cannot be made available to the public, data can be exported to eSIGN documents that can then be played back by an avatar (Hanke 2004).

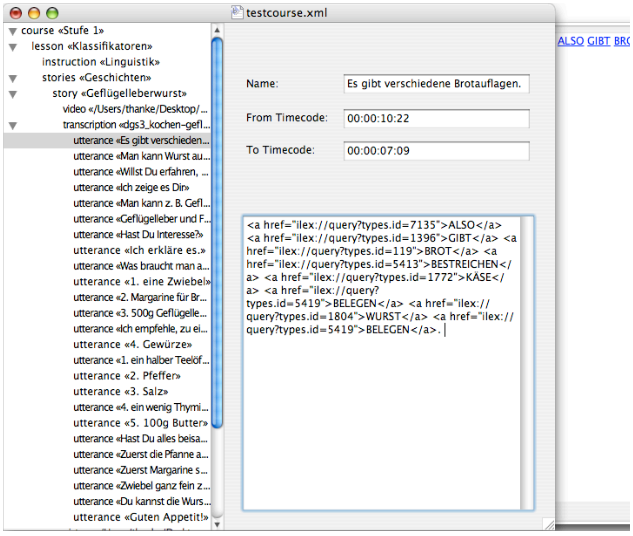
Importing data made available by other researchers as ELAN or SignStream documents requires a two-step approach. In a first step, data are imported into transcripts with only text tiers. In a second (optional) step, glosses as text should be replaced by database references. This step can only partially be automated, but finally results in transcripts that make full use of the iLex database structure.

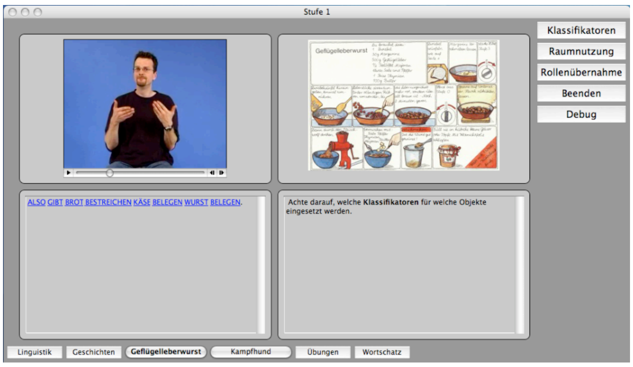
Whereas ELAN, SignStream, and iLex share the idea that tags label intervals of time and therefore can be thought of as variations of the concepts formalised by Bird and Liberman (2001), import from syncWRITER documents requires a number of assumptions as syncWRITER primarily tagged points in time. It can therefore become necessary to "repair" syncWRITER documents before or after the import process.

iLex supports the user in building metadata on all aspects of a signed discourse. For this, it supports all features required by Crasborn and Hanke (2003).

6. Applications in Teaching Materials

While we have produced high-quality sign language teaching CD-ROMs in the past (Metzger 2005), that have been individually programmed, we also see the need for less sophisticated, but easy and quickly to produce materials for our everyday teaching. Ideally, the lecturers should be able to do the complete production process themselves. Often the most complicated assets in e-learning materials for sign language is videos with time-aligned explanations and links, e.g. into a lexicon. The idea is to produce these assets as transcriptions in iLex, and then to import them into the authoring environment as complex content objects. We have therefore developed an authoring tool closely integrated with iLex. Through the interaction, links into a dictionary and the dictionary itself can be produced almost without any manual intervention. The player module, of course, works standalone and does not require a connection to the iLex database.





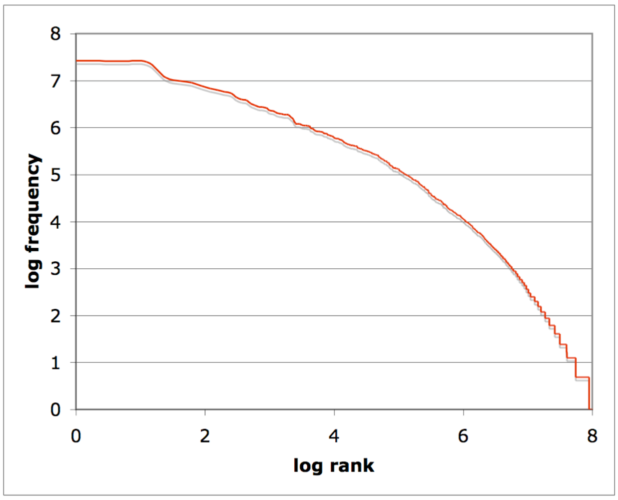
7. Zipf's Law for Sign Languages

When planning a general dictionary of DGS, there is no word list to start with. For a basic vocabulary, methods developed for spoken languages have been successfully adapted to result in a seed for a basic vocabulary of a signed language (Efthimiou/Katsoyannou 2001). For larger dictionaries, however, we see no alternative to a completely corpus-driven approach. The question then of course is how large a corpus needs to be in order to cover a sufficiently large portion of the lexicon.

For spoken languages, these predictions are often based on the rules of thumb referred to as Zipf's Law. The basic idea is that the product of the frequency of a word in a corpus and its rank is more or less constant over all words in the corpus.

Can we expect such a rule to also apply to sign languages? Function "words" play a significantly smaller role than e.g. in English, and it is not clear how productive signs fit into the game.

Certainly, we do not have a balanced corpus of reasonable size available to "verify" Zipf's law. Nevertheless, we did some math experiment with the current contents of the iLex database, only counting those tokens that refer to types undoubtedly qualifying as "lexical". These accounted for 108000 out of 125000 tokens. Surprisingly, the graph does look relatively smooth. While the graph is not exactly what you would expect for English, the low slope in the first ranks comes close to what Ha and Smith (2004) reported for Irish, a highly-inflected Indo-European language.



So we are tempted to "trust" Zipf predictions and use future work on the production of a general dictionary of German Sign Language to verify this.

8. References

Bird, S and M. Liberman (2001). A formal framework for linguistic annotation. *Speech Communication 33(1,2):* 131-162.

Crasborn, O. and T. Hanke (2003). Metadata for sign language corpora. Available online at: http://www.let.ru.nl/sign-lang/echo/docs/ECHO_Metadata_SL.pdf.

Crasborn, O., J. Mesch and E. v.d. Kooij (2004). *European cultural heritage online (ECHO): publishing sign language data on the internet.* Poster presented at TISLR 8 Barcelona, Spain.

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

Elliott, R. et al. (2004). An overview of the SiGML notation and SiGMLsigning software system. In: O. Streiter and C. Vettori (eds.): *Proceedings of the Workshop on Representing and Processing of Sign Languages, LREC 2004, Lisbon, Portugal*, pp. 98–104.

Ha, L. Q., and F. J. Smith (2004). Zipf and type-token rules for the English and Irish languages. In: *Proceedings of the Modelling for the Identification of Languages Workshop, Paris, France*, pp. 65-70.

Hanke, T. (2001). Sign language transcription with syncWRITER. *Sign Language and Linguistics*. 4(1/2): 275-283.

Hanke, T. (2002a). 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* (pp. 249-264). Hamburg: Signum.

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

Hanke, T. (2004). HamNoSys - Representing sign language data in language resources and language processing contexts. In: O. Streiter and C. Vettori (eds.): *Proceedings of the Workshop on Representing and Processing of Sign Languages, LREC 2004, Lisbon, Portugal*, pp. 1-6.

Hanke, T., R. Konrad and A. Schwarz (2001). GlossLexer – A multimedia lexical database for sign language dictionary compilation. *Sign Language and Linguistics* 4(1/2): 161–179.

Konrad, R. et al. (2004). *What's in a sign? Theoretical lessons from practical sign language lexicography.* Poster presented at TISLR 8 Barcelona, Spain.

Metzger, C. (2005). Die Firma. In: H. Leuninger and D. Happ (eds): *Gebärdensprachen: Struktur, Erwerb, Verwendung*, pp. 309-324. Hamburg: Buske.

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

Neidle, C. (2001). SignStream™: A database tool for research on visual-gestural language. *Sign Language and Linguistics*. 4(1/2): 203-214.

Pizzuto, E., and P. Pietrandrea (2001). The notation of signed texts: Open questions and indications for further research. *Sign Language and Linguistics* 4 (1/2): 29-45.

Prillwitz, S. et al. (1989). *HamNoSys. Version 2.0; Hamburg Notation System for Sign Languages. An introductory guide*. Hamburg: Signum.

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]

Translation of Natural Speech into Sign Language Based on Semantic Relations

Sandy Pleissner

Leipzig University
Faculty of Educational Science
Institute for Special Needs Education
Marschnerstrasse 29
04109 Leipzig, Germany
mail@sandypleissner.de

Abstract

Based on speech observations in children and categories of semantic classes we designed a system which identifies these in natural language and translates them into a sign language. To accomplish this translation, we use algorithms to annotate a set of semantic relations in children's language and hope to regain these sentences from natural source sentences. We define a set of rules used to change the word sequence of origin sentences at every marked relation.

1. Introduction

The paper describes the extraction of semantic relations in natural language based on syntactic and morphologic data we collect of language in children. This information exercises and improves the translation of language, here German, into sign language. Motivation is the assumption that semantic relations of the source language are more relevant than syntax for the arrangement of words and phrases in a target language for expressing equivalent information.

The syntax of sign languages differs considerably from the syntax of natural language sentences. It is often assumed that there are fewer, if any, grammar rules for sign languages. Each speaker follows his or her own grammar rules. The term "sign language syntax" is very speaker-dependent.

2. Processing

2.1. Child language

First, children's oral speech between birth and the first day at school is monitored (Stern and Stern, 1928). During development, the syntactic, morphologic and lexical abilities of the children are observed and thus the semantic categories the children are able to differentiate between is learned.

When the first steps to express wishes and feelings were observed, we recognized that the use of syntax rules are quite similar to sign language. At the age of seven, a child controls 15 base and main relations (Szagun, 1980). We isolated these relations and their typical syntacto-semantic realizations.

Two Examples with syntacto-semantic annotation:

Semantic Relation	*Examples*	*Annotation*
Handlungsrelation/ action relation	Er fährt Zug. *'he goes-by train'*	Subject, Object, Action – Verb
Lokative Handlungsrelation/ locative action relation	Er fährt nach Hause. *'he rides home'*	Subject, Object(Dativ), Action – Verb

2.2. Machine learning

In the second step a system is designed which focuses on the grammatical abilities of humans at seven years of age. A classifier is trained to extract the semantic relations, subject, object and direction of verb based on features such as part-of-speech verb type (Rudolph and Försterling, 1997), gender, case, tempus, numerus, mode and verb frames (Schulte, 2003). The lexicon is limited to the oral vocabulary of first to second grade children (Pregel and Rickheit, 1987).

Only non-complex sentences up to a length of 10 words and only one strong relation per sentence have been considered so far.

Semantic categories the automat must recognize are the three base relations presence, not-presence and again-presence. It's a necessary minimum to create new word sequences. Classification process is not bound to verb characteristics as action or state.

Semantic Relation	*Examples*
Vorhandensein/ presence	Oma ist am Fernseher. / Grandma is at the window. Papa liest ein Buch. / Dad reads a book.
Wieder-Vorhandensein/ again-presence	Ich will noch einmal. / I want again. Er fährt und fährt und fährt. / He drives and drives and drives.
Nicht-Vorhandensein*/ not-presence	Ich weiß nicht./ I don't know. Niemals!/ Never!

*Nicht-Vorhandensein/ not-presence implicates statements such as *denials* and *refusals*.

An more exactly assortment becomes possible, if we are able to split into another main relations.

The question– and intention–relation we separate to question- or w–question- relation and intention- and imperative–relation.

Examples:

Semantic Relation	*Examples*
Handlung/ action	Lisa rennt./ Lisa runs.
Lokative Handlung/ locative action	Lisa rennt nach Hause./ Lisa runs home.
Zustand/ state	Er ist rot./ He is red.
Lokativer Zustand/ locative state	Oma ist am Fenster./ Granny is at the window.
Intention/ intention	Ich will spielen./ I want to play.
Imperativ/ imperative	Lauf!/ Run!, Anne muss es tun./ Anne must do it.
Instrument/ instrument	Mit einem Messer schneiden./ Cut with a knife.
Dativ/ dative	Mama malt mit Anne am Computer./ Mum paints with Anne at the computer.
Handlung und Ort/ action and location	Ich lese im Buch./ I read in the book.
Bemerken/ recognize	Da sind Peter und Paul./ There are Peter and Paul.
Besitz/ Possession	Hans hat einen Hund./ Hans got's a dog.
Nachträgliche Bemerkung/ Addition	Oder es raschelt in der Wohnung./ Or it rattles on the roof.
Frage/ question	Seid ihr da?/ There are you?
W – Frage/ w-question	Wo seid ihr?/ Where are you?

For every relation we define one of these options:

- exists an explicit template for the target language
- exists a rule to insert one or more elements into another templates

As soon as we know contained relations, we start to identify contained constituents. Constituents are the atomically parts for rearrangement of sentences.

During analyzing sentences of child speech we noticed it's sufficiently to search for subject, verb, time, location, interjection, question word, negative and repeat particles. Another parts are quite rare, so we label these as objects.

We use some known techniques and tools To recognize constituents. Most of them are included in TNT–Tagger package.

2.2.1. Subject

We annotate sentences and marks contained subjects. Automat learns recognize subjects through tagging sentences, subjects and subject-frames (right and left neighbors). In later runs the automat is able to expand his knowledge to unknown sentences and sentence – structures.

2.2.2. Time constituents

The syntactic function of time constituents are similar to subjects. We can not define a set of preferred part-of-speech. Therefore we use again tags and frames of the annotated train corpus. Some words occurred more frequently in time frames. To support tagging we create a list of known time constituents.

2.2.3. Locative Constituents

Locative Constituents are often bound on appearing of prepositions. We observe prepositions of the train corpus and notice the neighbored words and tags.

2.2.4. Verb

Verb–recognition is a directly result of part-of-speech–Tagging. Usually verbs are labeled with standard tags or combinations.

Verb – Tag Examples	*Examples*
VVFIN	Gehst du jetzt?
VVIMP	Geh!
VVINF	Wir wollen gehen.
VVFIN + PTKVZ	Gehst du jetzt weg?
PTKZU + VVINF	Sie stand auf, um zu gehen.

2.2.5. Interjections

Interjections are marked too while part-of-speech – Tagging. Examples: Ah, oh, ieh.

2.2.6. Question words

Question words are question-introducing words beginning with character w. Therefore the relation is called w–question. The few existing question words the POS–Tagger recognize as interrogative pronouns. By taking the tagged pronouns at the beginning of sentences we got a list of question words.

Fragewort	*Examples*
was (Tag: PWS)/ what	Was tust du?/ What you are doing?
welche (Tag: PWAT)/ which	Welche Farbe ist das?/ Which color is it?

2.2.7. Negations

A negation becomes expressed only with a small number of words. Examples: Nein, kein, nicht/ no, not.

A negation is an indicator of an not–presence–relation.

2.2.8. Repeat

Similar to negations can we find particles of repeating, for example: schon wieder, noch einmal/ (already) again. Particles of repeat helps identifying again–presence–relations.

2.2.9. Object (or each other)

For the moment we are labeling all parts of speech as object, which was not tagged as Subject, time, locative, verb, interjection, question word, particle of negation or repeat.

2.3. Rearrangement of words

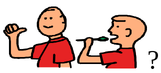
The third step is to find regularities in the order of the target sentence by considering German sign language standards and research (Prillwitz and Vollhaber, 1990) of the German sign language. The transformation rules are assigned to the semantic categories determined before. The rules of each category are stored in rules of a context sensitive grammar. During a translation the system analyzes a sentence, recognizes the relevant semantic relation, subject, object and direction of verb and yields a suggested translation after applying the transformation rules. This suggested translation fills the category template.

This rearranged sequence can be improved with the aid of an appropriate dictionary. Dynamic(al) dictionaries, which contain an animation, e.g. 3D animated DGS (Deutsche Gebärdensprache/German sign language), can use the suggested translation. For static dictionaries, e.g. PCS (Picture Communication Symbols/ Mayer-Johnson), sequences of subject and object are dependent on verb direction.

2.4. Addition

While processing we looked at simple cases of sentences. Some important (and necessary) parts of analyze are not discussed. Two short examples:

Two or more neighbored words are strong similar and therefore one or more are redundant.

INPUT	Ich esse ein Essen./ I eat a meal.
-> reducing number of words	Ich esse./ I eat. {ich}{essen}/ {i}{eat}
-> translate to target language -> PCS	 or  ?

Compounding New Items (Klima and Bellugi, 1979): Two or more words are part of a new or known sign.

INPUT	Ich öffne ein Fenster./ I open a window.
-> compounding	{ich}{fenster-öffnen}/ {i}{window-open}

INPUT	Ich esse Frühstücke./ I eat breakfast.
-> compounding	{ich}{frühstück-essen| frühstücke}/ {i}{breakfast-eat}

3. Examples

{*} means *dictionary entry*, (*) means *mutely*, *|* means *or*

3.1. Animated Dictionary (e.g. DGS)

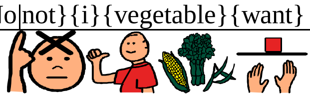
Semantic Relation	*Speech*	*Trans.Suggestion*
Handlungsrelation/action relation	Er fährt Zug.	{er} {zugfahren|Zug fahren}(.)
	Er fährt morgen Zug.	{morgen} {er} {zugfahren|Zug fahren}(.)
Frage-Relation/question relation	Wie ist dein Name?	{du} {Name} (ist) {wie}(?)

3.2. Symbol Dictionary (e.g. PCS)

Semantic Relation	*Speech*	*Trans.Suggestion*
Handlungsrelation/action relation	Ich rufe dich.	{ich} {du} {rufen}(.)[1]
	Ich lade dich ein.	{du} {ich} {einladen}(.)[2]

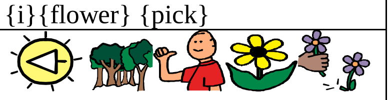
[1] *rufen:* left->right direction
[2] *einladen*: right->left direction

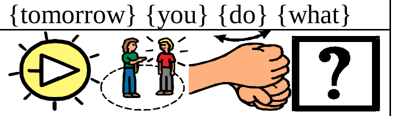
Example for minimum categorisation:

STEP	
INPUT	Ich will kein Gemüse./ I do not want vegetable.
-> identify Semantic Relation	Nicht-Vorhandensein/
-> analyze parts of sentence	Ich[Subject/ subject] will[Modaleverb/modal verb] kein[Negationspartikel/ particle of negation] vegetable.[Objekt/ object] .[]
-> choose template	{NEG} {SUBJECT} {OBJECT} {VERB}
-> insert base forms	{Nein|nicht|kein} {ich} {Gemüse} {wollen}/ {No|not}{i}{vegetable}{want}
-> translate to target language -> PCS	

Examples for optimum categorisation:

STEP	
INPUT	Ich habe gestern im Wald eine Blume gepflückt./ Yesterday I picked a flower in the forest.
-> identify Semantic Relation	Lokative Handlungsrelation/ locative action relation
-> analyze parts of sentence	Ich[Subject/ subject] habe gepflückt[Handlungsverb/action

	verb] gestern[Zeitkonstituente/ time constituent] im Wald[Ort/ location] eine Blume[Objekt/ object] .[]
-> choose template	{TIME} {LOCATION} {SUBJECT} {OBJECT} {VERB}
-> insert base forms	{gestern} {Wald} {ich} {Blume} {pflücken}/ {yesterday} {forest} {i}{flower} {pick}
-> translate to target language -> PCS	

STEP	
INPUT	Was machst du morgen?/ What do you do tomorrow?
-> identify Semantic Relation	Frage-Relation mit Handlung/ question relation with action
-> analyze parts of sentence	Was[Fragewort/ question word] machst[Handlungsverb/action verb] du[Subjekt/ subject] morgen[Zeitkonstituente/ time constituent] ?[Fragezeichen/ question mark]
-> choose template	{TIME} {SUBJECT} {VERB} {QUESTION WORD}
-> insert base forms	{morgen} {du} {machen} {was}/ {tomorrow} {you} {do} {what}
-> translate to target language -> PCS	

4. Conclusion

For the moment we can put all simple sentences into one of the existing templates as shown above. As a result translations of discussions between children are possible. Future questions we hope to answer which relate to adult speech are:

- How can we automatically create Meta templates from the described base relation?
- How can we merge the rule sets of two or more semantic relations?
- Can we split one sentence into many sentences with one relation?

5. References

Adam, Heidemarie (1993). Mit Gebärden und Bildsymbolen kommunizieren. Edition Bentheim Würzburg

Boyes Braem, Penny (1990). Einführung in die Gebärdensprache und ihre Erforschung. Signum Verlag.

Clahsen, Harald (1988). Normale und gestörte Kindersprache. John Benjamins Publishing Company Amsterdam/Philadelphia.

Klima, E. and Bellugi, U. (1979). The Signs of Language. Harvard University Press

Lindner, Gerhart (1997). Artikulation im technischen Zeitalter. Luchterhand.

Pregel, D. and Rickheit G.. (1987). Der Wortschatz im Grundschulalter. OLMS Verlag.

Prillwitz, S. and Vollhaber, Th. (1990). Gebärdensprache in Forschung und Praxis. Signum Verlag.

Rudolph, U. and Försterling, F. (1997). The Psychological Causality Implicit in Verbs: A Review. Ludwig-Maximilians-Universiät München.

Schulte im Walde, Sabine (2003). Experiments on the Automatic Induction of German Semantic Verb Classes. Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart.

Starcke, H. and Maisch, G.. (1977). Die Gebärden der Gehörlosen. Deutsche Gesellschaft zur Förderung der Hör-Sprach-Geschädigten e.V..

Stern, Clara und William (1928). Die Kindersprache. Verlag von Johann Ambrosius Barth.

Szagun, Gisela (1980). Sprachentwicklung beim Kind. BELTZ - Psychologie Verlags Union.

Articulatory Analysis of the Manual Parameters of the French Sign Language Conventional Signs

Annelies Braffort

LIMSI/CNRS, Orsay, France
Campus d'Orsay, Bat. 508, BP133, F-91403 Orsay cedex
annelies.braffort@limsi.fr

Abstract

This paper presents results of the analysis of French Sign Language (LSF) conventional signs that have been extracted from a LSF dictionary, in order to help the design of LSF processing systems. The signs (more than 1200) have been described, regarding manual parameters from an articulatory point of view. The movement parameter has been considered regarding the moving articulator: hand, wrist, and forearm. Thus, handshape, orientation and location parameters have been considered to be static or dynamic. The descriptions have been stored in a database, allowing us to compute quantitative data for each parameter and for the links between the parameters. Our analysis on this database gives us clues to design new description systems of lexical signs for SL processing, for automatic recognition or generation with the aim to design more accurate and synthetic representations.

1. INTRODUCTION

In most SL dictionaries (Moody, 1986), databases or other description systems, the lexical signs are described by means of four manual parameters, which are handshape, location, orientation and movement, and facial expressions, where each parameter is systematically specified, in a uniform way.

This is generally not suitable for processing systems dedicated to SL, neither for analysis by means of image processing (Bowden et al., 2004 ; Lenseigne & Dalle, 2005), nor for automatic recognition (Braffort, 1996) or generation (Hanke, 2000; Huenerfauth, 2004). Most of these systems need to integrate more precision on each parameter and on their inter-relations.

To help the design of such systems, we have conducted a study on LSF lexical signs, by describing the pictures contained in the first volume of a LSF dictionary (Moody, 1986) one by one, that is to say 1257 signs.

Each sign has been described according to visual features. In order to be able to obtain quantified information, descriptions have been stored in a data base (OpenOffice Base). The exploitation of the database enables us to obtain quantified information on the frequency of the values of each feature and the relation between values of various features.

This paper describes the chosen point of view while defining the parameters, especially regarding the movement parameter, and gives some results on the handshape, location and orientation parameters.

It describes the methodology and sketches out the potential observations that can be provided by this kind of mixed annotation.

2. DEFINITION OF MOVEMENT

The muscles which make it possible to move the fingers are different from those which move the hand around the wrist axes, and are different from those which make it possible to move the arm. Thus, differently to what is considered in linguistics studies, we burst the description of the movement in the description of the three parameters of handshape, orientation and location. The handshape can be static or dynamic, just as the orientation and the location. That makes it possible to account for the articulatory phenomena better, and especially the fundamental difference which exists between the movement and the other manual parameters: From an articulatory point of view, there are always a configuration, an orientation and a location of the hand, but not always a movement. These three parameters are described with visual criteria, from the point of view of the signer.

For example, we describe the sign $[\text{WORM}]_{\text{LSF}}$ (Figure 1), in the following way (only the active hand is described):

- The handshape is dynamic, the index has an wiggling movement, alternating from an extended position to a semi-bent position, the other fingers being bent.
- The location varies in a neutral zone located in front of the signer. The trajectory is linear, in a horizontal frontal axis, and the movement is directed towards the left.
- The orientation is static: The hand axis is directed towards the left, the palm is directed to the bottom and the wrist is not bent;

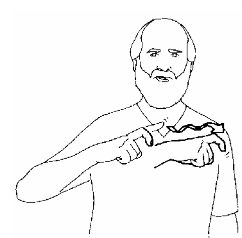


Figure 1: $[\text{WORM}]_{\text{LSF}}$ (Moody, 1986)

Thus, the specificity of this description is that the index movement is dissociated from the arm movement, unlike the schematic description given in the dictionary, in which the movement is represented by only one symbol: A horizontal wavelet, concatenation of the index and arm movements.

This paper presents results of an analysis on handshape, location, orientation and on more global aspects.

With such a corpus, linguists and computer scientists can study the same video together, so as to perform complementary analysis.

3. CONFIGURATION

The study has shown that there is a great diversity of handshapes (139 different configurations), static or dynamic. We indexed them and calculated the percentages of occurrence for each of them in the dictionary. Only 65 handshapes appear more than twice. Table 1 lists the most frequent handshapes and their occurrences.

Out of the 1257 signs, 62.5% are bimanual and 37.5% are monomanual.

A finer study was undertaken on the handshape occurrences of the strong hand according to whether the sign is bimanual (index, mitt, s, flat, 5, v, angle, 1, beak, key, clip, x, ball, y) or not (index, flat, 1, v, y, s). When the sign is bimanual, we measured the handshape occurrences of the weak hand (mitt, flat, s, index, 5, angle, 1, beak, ball). For more of 75% of the bimanual signs, the two handshapes are identical (index, mitt, 5, flat, s, angle, 1, ball). When the handshapes are different, the handshape of the weak hand is simple (mitt, flat, s, index).

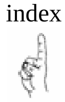
Name	index	flat	mitt	s	5
					
Occur	122	78	71	69	59
.Name	1	v	angle0	y	beak
					
Occur	52	52	40	38	36
Name	key	ball	clip	x	u
					
Occur.	35	31	30	30	24
Name	bent2	c	hook	bmiddle	5p/beak
					
Occur	23	22	22	20	18

Table 1: Most frequent handshapes in LSF.

The dynamic handshapes raised in the dictionary are handshapes for which at least one of the finger joint undergoes a variation during the sign. Nearly 20% of the signs present a dynamic handshape, corresponding to 12% (152) of closing handshape (at least a finger joint is inflected during the sign) and 6% (75) of opening handshape (at least a finger joint undergoes an extension during the sign). More details on closing and opening handshapes can be found in (Braffort, 1996). Less than 3 % of the signs present more complex finger movements from an articulatory point of view. All in all, one can say that the more complex the dynamic aspect is, the more the handshape is simple.

Generally, we can note that the articulatory realization of a handshape always remains in an acceptable level of complexity: There exists few dynamic handshapes and they consist of an initial and a final simple handshape. The most frequent handshapes seem to correspond to the proforms used in LSF (Cuxac, 2000). This last property must be validated with statistics on LSF corpora from which the proforms are annotated.

Regarding these tables, representations of handshape in automatic recognition systems can be simplified and optimized, like in (Braffort, 1996), in order to enhance discrimination between the different handshapes.

4. LOCATION

Hand location can be described on two levels of granularity: First of all on a *global level*, in order to distinguish the zone in which the sign is carried out, and then, on a *finer level*, to analyze the variation of the location within this zone, when the arm undergoes a movement.

Location values correspond to zones in space, more or less wide. For most of the studied signs (more than 60%), hand location is in the neutral zone placed in a half-sphere in front of the signer. We found 48 different locations, 14 corresponding to more than 92% of the signs. In 96% of the cases, even if we observe a motion of the arm, the signs are carried out within only one zone.

Describing the trajectory of the wrist in space specifies motions. In order to differentiate all the types of motions, their trajectories are defined compared to the various axes and planes of a 3d coordinate system centred on the signer. It is illustrated in Figure 2.

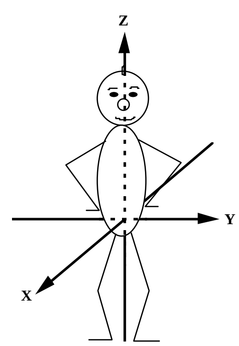


Figure 2: The 3d coordinate system centered on the signer.

Some results are presented in a synthetic way here. More details can be found in (Braffort, 1996).

More than 40% of the signs trace a straight trajectory; 82% of these signs are parallel to one of the three axes, and nearly 15 % are parallel to a plane formed by two of the axes. Nearly 88% of the signs present a curved trajectory included one of the three main planes. Nearly 92% of the signs whose trajectory is a circle are parallel to one of the three main planes.

In vertical movements, the signs for which only one arm is moving are frequent. For the lateral movements, the signs for which two arms are moving with opposite motions are frequent. For the horizontal and sagittal movements, the signs for which two arms are moving in a parallel way are frequent. Moreover, in a general way, when the two arms are moving, they always have symmetrical movements (parallel, opposed or shifted).

The global result is that the simplest movements to realize are also the most frequent, which goes in the same direction as the observations concerning the handshape parameter.

This analysis also gives us clues for the definition of macro representation of one-handed or two-handed arm motion primitives for automatic generation.

5. ORIENTATION

The description of the hand orientation is very tricky. It is often described in an absolute way, compared to the 3d coordinate system presented above, such as in HamNoSys (Prillwitz & Leven, 1989). The orientation of the hand is defined in this type of system by three items:

- The state of the wrist (extended or bent, at rest or in rotation),
- The direction of the hand axis,
- The direction of the hand palm.

This system is well adapted to describe signs such as [TABLE]$_{LSF}$ or [CEILING]$_{LSF}$ (Figure 3) because these signs draw in space salient iconic features for which absolute and static orientation of the hands is necessary.

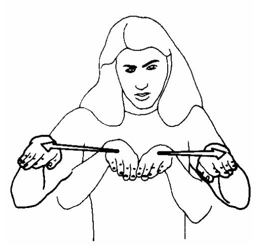


Figure 3: [TABLE]$_{LSF}$ (Moody, 1986)

Figure 3: [CEILING]$_{LSF}$ (Moody, 1986)

Compared to the 3d coordinate system, the hand orientation can vary during the sign if at least one of the directions varies or if the wrist undergoes a motion around one of its joint axes. For gestures for which the orientation undergoes a variation during the sign, such as [CORRIDOR]$_{LSF}$ (figure 4), the use of an absolute coordinate system is not easy.

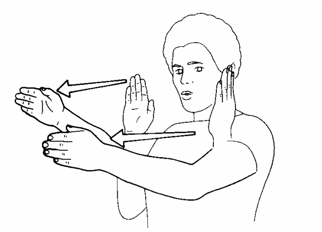


Figure 4: [CORRIDOR]$_{LSF}$ (Moody, 1986)

For these signs, a system based on a relative coordinate system fixed on the forearm of the signer and specifying the state of the elbow (flexion, extension, rotation) would be perhaps more suitable.

Moreover, for most of the studied signs, hand orientation seems to be a consequence of the various joint movements, rather than an intentionally selected orientation in the signer coordinate system.

The description of the hand orientation remains an open question that does not have a satisfying solution at the present time, synthetic enough for an implementation in a processing system.

6. INTERDEPENDENCES

The database is also used to study the relations and interdependences between the various parameters:

- Relations between the handshape and the movement: In the bimanual signs, we observed that when the two arms are moving, the two handshape are identical, while when only the arm of the strong hand is moving the two handshapes are different.
- Relation between the movement and the orientation: According to the type of hand trajectory, the behaviour of the orientation is different. For linear or circular trajectories, the orientation most of the time is static. On the other hand, for arched trajectories or when the hand does not move, the orientation is rather dynamic.

Many other results can be obtained from this database. For example, it is possible to observe the correlation between the complexity of the handshape and the complexity of the arm motion. As one could expect it, the more the arm motion is complicated, the more the handshape is simple and reciprocally.

Here also, the analysis of the database can give us clues for the definition of macro representation of gestural units combining two or more articulators for automatic generation and recognition.

7. Conclusion

From the study we have conducted on 1257 lexical sign in a LSF dictionary, we have obtained several kinds of result.

The first one is related to the status of the classical four manual parameters. In the context of dictionaries, if handshape and location can be described by giving a value chosen in a predefined list, even considering the dynamic side of these parameters, this is not the case of the orientation. That means that this parameter cannot be considered at the same level than the others and much more work must be done on this topic.

The second kind is related to the occurrences of the values for each parameter. This allows us to design our processing tools in an iterative way, beginning to represent the most frequent phenomena, and to design more discriminante representations.

The third one is related to the interdependences between the parameters. The properties of interdependency that we can extract from the database can be of great help in the design of constraints, allowing us to design a simpler description of signs in a processing system by the way of macro representations combining two or more articulators.

Finally, this study give us the feeling that new studies should be achieved on the description of lexical signs, allowing a more accurate and synthetic representation for SL processing, such as in (Filhol, 2006).

This database is being extended to allow more descriptions on other articulatory phenomena, regarding the elbow and the shoulder, and more descriptions about the iconic intent in the realization of signs.

8. References

Bowden R, Windridge D, Kadir T, et al.. (2004). A Linguistic Feature Vector for the Visual Interpretation of Sign Language, In Pajdla, Matas (Eds), Proc. 8th European Conference on Computer Vision, ECCV04. LNCS3022, Springer-Verlag, Volume 1, pp391- 401.

Braffort, A. (1996). Reconnaissance et compréhension de gestes, applicationà la langue des signes. Thèse de Doctorat en Informatique (PhD thesis). Université Paris-sud, Centre d'Orsay.

Cuxac, C. (2000). La Langue des Signes Française (LSF). Les voies de l'iconicité. Faits de Langues n°15-16, Ophrys.

Filhol M. & Braffort A. (2006). A sequential approach for Sign Description. LREC 2006, Workshop proceedings : Representation and processing of sign languages (in these proceedings).

Hanke, T. (2004). HamNoSys for representing sign language data in language resources and language processing contexts. In: Streiter, Oliver / Vettori, Chiara (eds): LREC 2004, Workshop proceedings: Representation and processing of sign languages. Paris : ELRA - pp. 1-6

Huenerfauth, M. (2004). Spatial Representation of Classifier Predicates for Machine Translation into American Sign Language. In: Workshop on the Representation and Processing of Signed Languages, LREC 2004, Portugal.

Lenseigne, B. and Dalle, P. (2005). A tool for Sign Language Analysis through Signing Space representation. In: Sign Language Linguistics and the Application of Information Technology to Sign Languages, Milan, Italie

Moody, B. Moody B. (1986). La langue des signes. Dictionnaire bilingue élémentaire. vol.2, IVT, Paris.

Combining video and numeric data in the analysis of sign languages within the ELAN annotation software

Onno Crasborn[1], Han Sloetjes[2], Eric Auer[2] & Peter Wittenburg[2]

[1] Department of Linguistics, Radboud University Nijmegen, [2] Max Planck Institute for Psycholinguistics, Nijmegen

Abstract

This paper describes hardware and software that can be used for the phonetic study of sign languages. The field of sign language phonetics is characterised, and the hardware that is currently in use is described. The paper focuses on the software that was developed to enable the recording of finger and hand movement data, and the additions to the ELAN annotation software that facilitate the further visualisation and analysis of the data.

1. Introduction

The ELAN software is a linguistic annotation tool that was designed for the creation of text annotations to audio and video files. Starting its life in the domain of speech and gesture research, in recent years it has increasingly been used in the study of sign languages. Aspects of ELAN were enhanced for the creation of the sign language corpus in the EU project European Cultural Heritage Online (ECHO), in which comparable data from three signed languages were collected, annotated, and published (Brugman et al. 2004, Crasborn et al. 2004).

This paper reports on a recent development in ELAN: the integrated display and annotation of numeric data with audio and video data. These numeric data presently stem from one specific movement tracking system that focuses on the hand and fingers, but in principle any type of numeric data can be used. In this paper we focus on the use of such technology in the phonetic analysis of sign languages used by Deaf communities (section 2). We describe the current hardware that is used, as well as the specific nature of the kinematic data generated by the software that we developed (section 3). Section 4 starts with a brief sketch of the ELAN annotation software, and describes in detail what types of functionality have been added to integrate the display and analysis of the kinematic data. We conclude by presenting some possible future software developments that would broaden the use of the current enhancements in ELAN, and discuss ways in which the integration of video and kinematic data can be used for all kinds of exploratory studies in sign linguistics (section 5). However, the focus of the present paper is not on concrete phonetic questions that might be investigated in sign languages, but rather on the methodology of such studies.

2. Sign language phonetics

Just as in spoken languages, sign languages show a distinction between a phonological and a phonetic level of organisation. Lexical items in sign languages typically consist of manual action: short one-handed or two-handed movements. These signs can be characterised in terms of properties like handshape, orientation, location and movement. The lexicon of a sign language consists of a limited number of recurring handshapes, locations, etc. These forms make up the phonological inventory of the language, and have been compared to phonemes or phonological features since Stokoe's (1960) pioneering work (see Brentari 1998 for an overview). In addition, non-manual aspects play a large role in sign language structure, but primarily at the morphosyntactic, pragmatic and discourse levels, and to a much lower degree in the lexicon (see Crasborn 2006 for an overview).

While these phonological specifications are constant for a given lexical item, the phonetic realisation of a sign by a given person at a given moment is highly variable. This phonetic variation has received fairly little attention in the literature (but see Wilbur & Nolen 1986, Wilbur 1990, Wilcox 1992, Cheek 2001, Crasborn 2001). In part, this has been due to the lack of easily accessible tools for measuring the articulation of sign language – the most obvious and accessible level of structure to quantify, and in that respect similar to acoustics in speech. Studies of the phonetic level of sign language could contribute considerably to our linguistic and sociolinguistic knowledge of sign language structure and use.

Phonetic studies that are based on video are limited in their quantitative potential: they are typically used as the basis for transcription, and are restricted in having a fairly low temporal resolution (25Hz for PAL and 29.97Hz for NTSC). The equipment that is described in section 3 allows for the very detailed measurement of hand and finger movements (and in principle any body parts), both in space and time. The aim of the software development described in this paper is to facilitate phonetic studies. The acquisition and use of equipment will remain problematic given the high costs, but the data analysis will become much easier with the extensions of ELAN described in section 4: the user will be able to integrate the measurements with video recordings, which will facilitate data analysis for phonetic studies.

3. Current equipment for measuring hand and finger movements

3.1. Hardware

The lab setup includes one right hand Virtual Technologies *CyberGlove* with 22 bend sensors and two Ascension *Flock of Bird* electromagnetic location trackers. The glove also has a button and a LED light on it. Those can be used for generic user interface purposes or to place or visualize time markers. One tracker is fixed to the glove, the other tracker can be attached to the other hand or to other reference points on the subject.

The CyberGlove equipment consists of a glove with proprietary resistive bend sensors and a box with electronics. The box is connected to the PC via a RS232 serial port. The two location trackers work with one common transmitter antenna cube. Each tracker has its own electronics, communicating with the other through a fast serial RS485 link. Only one of the tracker electronics boxes is connected to the PC via the second RS232 serial port of the PC.

The glove only measures spreading of fingers, not the absolute sideways finger movement or position. This should not be a problem when analyzing gestures. The wrist joint has a large bend radius, so the wrist flex and wrist abduct sensors need relatively careful calibration. Most other sensors measure the bending of joints with a small bend radius. As long as the sensor is long compared to the bend area, little calibration is needed. Finally, the thumb movement is so complex that it is hard to capture in terms of few bend angles. Even with good calibration, the relative position of the thumb with respect to the fingers will not fit the real hand shape very well. This has to be taken into account for sign language analysis.

The glove electronics use analogue low-pass filters to attenuate the spectral components above 30 Hz at 20 db per decade. The documentation tells that human finger motion has been found to be usually slower than 3 Hz (that is, three movement cycles per second). The default sampling rate is 90 Hz, which is close to the maximum. The limiting factors are the analogue to digital conversion time and the serial data transfer. The high sampling rate is assumed to be useful to pinpoint the onset time of a movement. Sensor values will be integers in the range of 40 to 220 for most hands. For most sensors this means a resolution of about two units per degree.

The tracker electronics have a default setting of 104 Hz sampling rate. Higher rates (up to 144 Hz) are possible but not recommended. Both location and orientation are measured. Measures are represented as signed 14 bit integers which, at the default 91 cm range coordinate system, gives a resolution of roughly 1 mm. Angular resolution is about 0.1 degrees. The accuracy is about 0.5 degrees / 2.5 mm RMS averaged over the translational range.

3.2. Software

The trackers come with a text based configuration tool called CBIRD. For the glove and for the kinematics model of the hand, we use the Immersion Virtual Hand toolkit. As the Virtual Hand software does not support multiple tracker setups well, the CBIRD software has to be used to configure the Flock of Birds hardware manually before the Virtual Hand software can use both trackers simultaneously. The Immersion software consists of a Device Manager that manages hardware connections, the DCU configuration tool and a library (which comes with some sample programs).

Defaults are not well suited for use with two trackers, so both CBIRD and DCU have to be used to get the proper configuration after booting the hardware. In addition, the communication protocols are well documented. The advantage in using the Virtual Hand toolkit is that it contains a kinematic model of the hand geometry and movement abilities and a (mostly undocumented) interface to that.

The GloveGUI software simultaneously logs and visualizes the data from the trackers and the glove. Several aspects of the logging (like precision and speed) can be configured through command line options. The effective data rate is limited by the polling speed of Device Manager, as GloveGUI fetches the measurements from Device Manager. The visualization in turn fetches coordinates from the logged data. Default rates are 20 window updates and 25 log samples per second, but one will usually use higher log rates for sign language analysis.

The graphical user interface of GloveGUI is designed to minimize interaction: The user has to decide about log file processing (overwrite, append, abort) and has to decide whether to proceed if not all data sources are ready. Normally, the user only has to confirm the current setup to start the recording. To end the recording, it is enough to close the main window or hit the ESC key.

During recording, the three mouse buttons can be used to toggle the usage of tracker location information, to centre the hand display, and to cycle through the available viewing directions.

The log files are plain text, and the first sample of each log file is annotated with comment lines. Each sample consists of the following data: PC timestamp, glove electronics timestamp, switch/LED state, the 22 glove sensor bend angles (3 for each finger, 2 for the wrist, further sensors for the spread between fingers etc), tracker data (3D location, 3D direction). The 3D directions are available as a vector and as a triple of azimuth, elevation and roll. The following derived values are available: wrist location and rotation, end points of each finger bone (20 coordinate triples), and the movement distance for each point compared to the previous sample. The finger bone end point logging can be disabled with a command line option of GloveGUI.

4. The extension of ELAN to cover the analysis of numeric data

4.1. The ELAN annotation tool

ELAN is a linguistic annotation tool for digital video and or audio, which is available for several operating systems: Linux, Windows, and MacOS X. It provides integrated video/audio playback and supports the association of up to four video files with the annotation document. An unlimited number of annotation layers or tiers can be created. Typically, a tier groups annotations that describe the same phenomenon and that share the same constraints on structure and or content.

Two main concepts in ELAN's design are "media players" and "viewers".

ELAN creates a media player for each associated audio/video file. If media files are out of sync, they can be synchronized to each other within ELAN. ELAN provides a rich set of player controls to navigate through the media. The smallest step unit is 1 millisecond, which means that the media play head can be positioned with a maximum precision of 1 ms. Consequently, the boundaries of an annotation can be determined with millisecond precision.

The sample frequency of video is in most cases 25 or 30 frames per seconds, resulting in a frame duration of 40 or 33 ms. For these type of media, ELAN enables the user to step through the media frame by frame. Here the millisecond precision would not add anything. The higher frequency of the kinematic recordings described in section 3 better exploits ELAN's precision. As was described in the preceding section, kinematic recordings

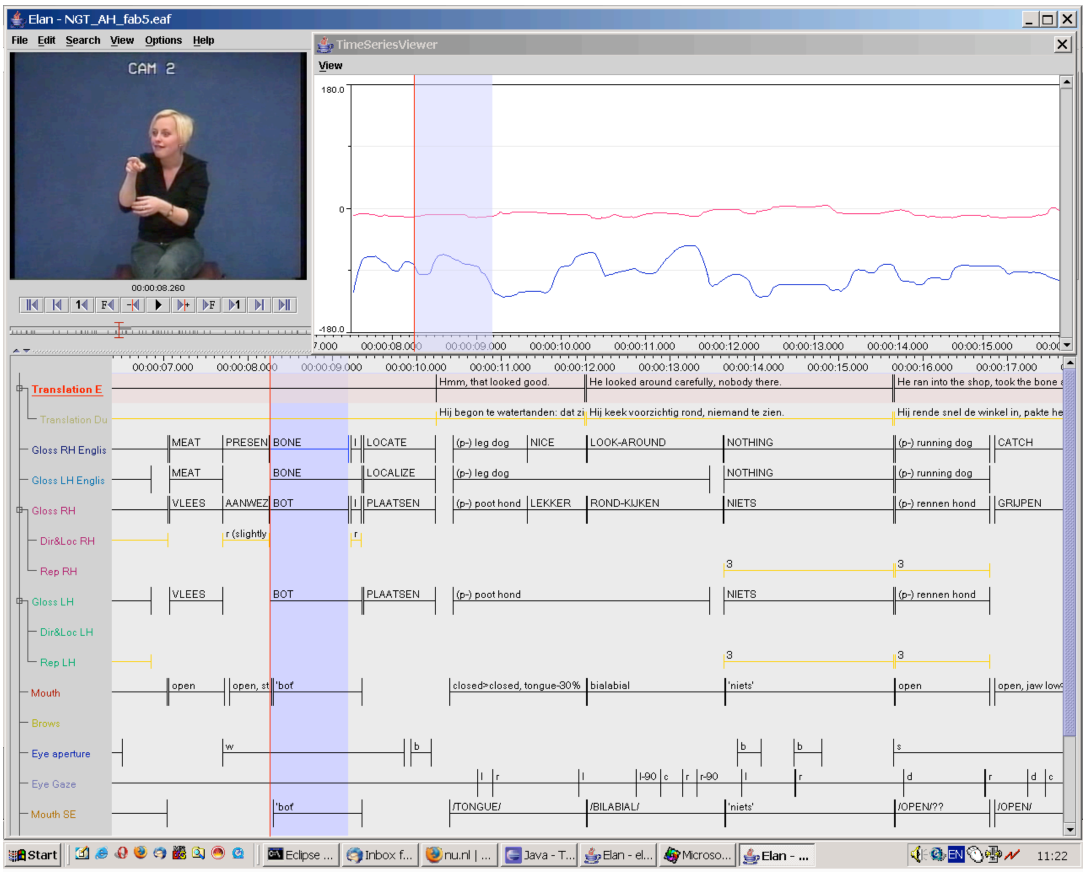


Figure 1: ELAN's media (top left), time series (top right), and timeline (bottom) viewers.

typically contain data with a frequency of around 100 Hz. This implies a significant increase in temporal precision for researchers studying sign languages. At this moment, ELAN offers no controls to navigate through the media in steps of 5 or 10 milliseconds: the user has to use the millisecond controls.

ELAN can display several viewers (see Figure 1), most of which present the underlying annotation data in their own specific way. However, there is also a viewer that can display a waveform of an audio file. All viewers are synchronized, i.e. they are connected to the media time and they all have some way of indicating the current media time and the selected time interval.

4.2. Integration of CyberGlove data in ELAN

A CyberGlove data file is associated with an annotation document in the same way as an audio or video file. Instead of a player, a specialized viewer is created to visualize the content extracted from the CyberGlove samples.

This viewer can contain multiple panes, and each pane can contain multiple movement tracks; the user is able to configure these.

4.2.1. A specialized reader for the proprietary CyberGlove file format

The CyberGlove software produces samples that consist of 40 - 100 distinguishable measurement values (see the description in section 3). A specialized reader or import module that is aware of the structure of such file has been developed, which is capable of dealing with the variations that can occur in this kind of files and is able to

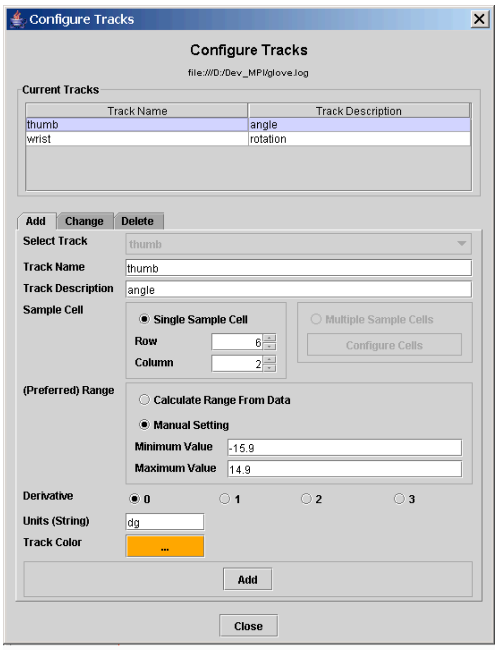


Figure 2: The track configuration pane.

calculate the sample frequency from the timestamps in the samples: the sample frequency is not explicitly listed in the data file. For each field in the samples, the reader is able to create a track. A track consists of an array of (single) measurement values and some visualization attributes.

4.2.2. Facility for track selection and calculation of derivatives

Given the multitude of measurement fields per sample, it is not feasible to simply visualize all information captured in the samples. Users need the opportunity to compose their own selection of tracks, based on the particular interest at hand. Therefore, a user interface has been developed to enable the selection and customization of tracks (Figures 2 and 3). Often the interest of the researcher goes beyond the bare data available in the file. For example, the amount of change over time in a certain measurement field could be equally important as the measurement itself, leading to the need for derived tracks. From a track recording the movement (covered distance) of a certain point on the glove, tracks for velocity, acceleration and jerk can be derived. For jerk, an extra filtering step might be necessary to reduce the noise in this type of derived data.

4.2.3. Facility for synchronization of the glove data to the media file(s).

Since it is very unlikely that the video files and CyberGlove recordings start at exactly the same time, the two signals need to be synchronized in some way. For that reason, the synchronization facility that already had been implemented in ELAN for audio and video files has been extended to support synchronization of video and CyberGlove streams. Corresponding events in both streams can be identified based on the graphical representation of the CyberGlove data. Thus, a new time origin for one or both streams can be determined and stored, guaranteeing synchronous playback in ELAN. To facilitate this process, use can be made of the button on the CyberGlove that switches a LED on and off: this will be visible in the video recordings, and changes the value of one of the parameters in the log file from 0 to 1 or the other way round (which can be visualised in one of the tracks).

4.2.4. Synchronized/connected viewer for CyberGlove data (time series viewer)

The tracks extracted from the CyberGlove file can be visualized as line plots, parallel to a (horizontal) time axis. A new viewer had to be created for this kind of data, a Timeseries viewer. The Timeseries viewer has a

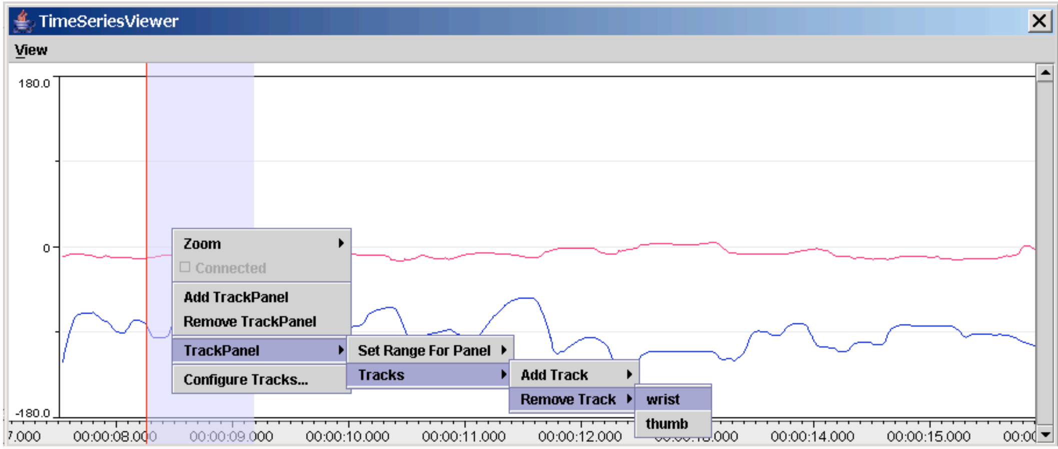


Figure 3: A pop-up menu to adjust what is displayed in the time series viewer.

horizontal time ruler, like the ELAN Timeline- and Waveform Viewer. The horizontal resolution can be changed by zooming in or out, effectively changing the number of milliseconds each pixel on the screen represents.

In addition to a horizontal ruler the time series viewer has also a, simple, vertical ruler labeling the minimum (low value) and maximum (high value) of the viewer area and a few values in between. A requirement specific to this viewer comes from the variety of value ranges that tracks can have. A track holding rotation values in degrees might have a range from -180 to 180 (or from 0 to 360) while a track representing velocity values might have a range from 0 to 20 and yet another track might have values between 0 and 1. The usability of the visual representation of the tracks would vanish when all these tracks would have to share the same coordinate system. To avoid this, the viewer can have multiple areas or ‘panels’, each one of which can display multiple tracks that can reasonably share the same range of amplitudes (coordinate system). Tracks can be added to and removed from a panel, panels can be added to or removed from the viewer (see Figure 2).

All viewers in ELAN have some way to visually mark the selected interval and the crosshair position (current media time). The Timeseries viewer marks these entities in the same way as other viewers with a time ruler, the Timeline viewer and the Waveform viewer (a light blue rectangle for the selected interval and a vertical red line for the crosshair).

In the Timeseries viewer the selected interval can be used for the extraction of the minimum or maximum value within that interval on a certain track.

4.3. Integration of glove data with annotations

The benefit of the integrated visualization of numeric data with video and annotations is twofold: on the one hand the line plots can be assistive in accurate annotation creation, on the other hand numeric values can be transferred to annotation values. The latter feature is very important for quantitative research.

This transfer has been implemented as follows: for each annotation on tier X, extract minimum or maximum in the annotation’s interval from track Y and add the result to a (new) child annotation on tier Z (being a child tier of tier X). The resulting annotations can be included in the export to tab-delimited text for further processing in a spreadsheet or in an application for statistical analysis.

5. Future developments and use

5.1. Inclusion of other data formats

The initial efforts in the area of integration of numeric data in ELAN have been geared towards full support for the very specific data files generated by the *Cyberglove* data glove and *Flock of Birds* motion sensors. However, in the design and implementation stage, a much broader application has always been in mind.

The components developed for the CyberGlove data are suited for any kind of ‘time series’ data, i.e. any data consisting of a list or array of time - value pairs, produced by whichever device or software application. Eye tracking equipment is just one example.

The main problem here is that there does not seem to be a standard format for such data. It is therefore unavoidable to write special ‘readers’ or import modules for each kind of file. Such modules have been created for the CyberGlove files and for a very simple, generic time - value file, a text file where each line holds a timestamp and a value, separated by a white space.

A Service Provider interface has been defined to enable addition of other, custom import modules. Third parties should be able to create and add their own module to the software independently from ELAN release dates.

5.2. Addition of phonetic data from spoken languages

ELAN's new functionality extension to represent and present time series data can also be seen as an opening to include phonetic data from spoken languages. For a number of research programs it becomes increasingly interesting to combine different data streams for analysis. In phonetics, it is interesting to combine the speech wave with video recordings of lip movements and other visible articulators, and with laryngograph and articulograph recordings, for example. ELAN makes it possible to easily synchronise these streams and present them in a synchronous presentation together with layers of annotation. In this respect ELAN offers unique functionality.

5.3. Linguistic uses

For the type of phonetic research on sign languages characterised in section 2, the present enhancement of ELAN offers two separate types of benefits. Firstly, at a practical level, the visualisation of the articulatory (kinematic) recordings in parallel with the video recordings allows for very efficient segmentation of these kinematic data: the test items can be quickly identified in the video window, and then precisely segmented using the higher temporal resolution of the movement tracks.

Secondly, in addition to the targeted recording and analysis of experimental phonetic data, a quite different type of use is also envisaged. Linguists working on phonological and prosodic aspects of sign languages can use the visualisation of the arm, hand and finger movements to generate research questions and hypotheses about various aspects of the form of signing. It has been difficult to develop sensible notions about movement features such as size and 'manner' (tenseness, speed), as well as 'holds' (pauses), without a common view on what exactly these properties refer to. While the phonological categories need not necessarily be expressed in terms of kinematic features, the ability to explore sets of video data with a view on displacement, velocity and acceleration of selected points on the hand is foreseen to have great benefits and lead to new insights on the form of connected signing.

In this respect, a great advantage of the present approach to accessing and processing data from movement trackers is that the data are stored and integrated in a multimedia corpus, rather than being data that are only used for an experiment and then disregarded. Metadata that characterise the overall properties of the recording situation (as in the IMDI standard, for example; see Broeder & Offenga 2004) are flexible enough to include a description of the general properties of the kinematic recordings, including a reference to the data file. In this way, the wider use of the data beyond the experiment at hand is indeed a feasible option.

6. References

Brentari, D. (1998) *A prosodic model of sign language phonology*. Cambridge, MA: MIT Press.

Broeder, D. & F. Offenga (2004) IMDI metadata set 3.0. *Language Archive Newsletter*, 1-2: 3.

Brugman, H., O. Crasborn & A. Russell (2004) Collaborative annotation of sign language data with peer-to-peer technology. In: *Proceedings of LREC 2004, Fourth International Conference on Language Resources and Evaluation* (pp. 213-216). M.T. Lino et al., eds.

Cheek, A. (2001) *The phonetics and phonology of handshape in American Sign Language.* PhD thesis, UT Austin.

Crasborn, O. (2001) *Phonetic implementation of phonological categories in Sign Language of the Netherlands*. Utrecht: LOT. (PhD thesis, Leiden University.)

Crasborn, O. (2006) Nonmanual structures in sign languages. In: *Encyclopedia of Language and Linguistics, 2nd ed.* (pp. 668-672) K. Brown, ed. Oxford: Elsevier.

Crasborn, O., E. van der Kooij, D. Broeder & H. Brugman (2004) Sharing sign language corpora online: proposals for transcription and metadata categories. In: *Proceedings of the LREC 2004 Satellite Workshop on Representation and processing of sign languages* (pp. 20-23) O. Streiter & C. Vettori, eds.

Stokoe, W. (1960) *Sign language structure. An outline of the visual communication systems of the American Deaf.* Silver Spring, MD: Linstok Press.

Wilbur, R. (1990) An experimental investigation of stressed sig production. *International Journal of Sign Linguistics* 1: 41-59.

Wilbur, R. & S.B. Nolen (1986) The duration of syllables in American Sign Language. *Language and Speech* 29: 263-280.

Wilcox, S. (1992) *The phonetics of fingerspelling*. Amsterdam: John Benjamins.

Forming Sign Language Learning Environments in Cyberspace

Maria Mertzani, Clark Denmark, Linda Day
Centre for Deaf Studies, University of Bristol
8 Woodland Rd, BS8 1TN, Bristol, U.K.
E-mail: M.Mertzani@bristol.ac.uk, A.C.Denmark@bristol.ac.uk, Linda.Day@bristol.ac.uk

Abstract

In this paper we would like to present the way virtual learning environments (VLEs) are employed into the teaching and learning of British Sign Language (BSL) at the Centre for Deaf Studies of Bristol University, U.K. By considering cyberspace a culturally constructed environment where people can form different virtual communities, this paper will focus on the creation of a virtual learning community for the purposes of BSL learning. Both tutors and students have access and meet on two main websites: SignStation and DeafStation, from where they can retrieve authentic BSL material during their classes and interact through a videoconferencing software system, Panda. We describe the development of VLE and discuss the practices employed when meeting online in terms of instruction delivery and knowledge construction.

1. INTRODUCTION

The World Wide Web is an emerging technology that facilitates and supports the process of learning and teaching a second language (L2) since it is a tool for accessing to learning resources as well as for communicating from distance. It integrates interactive media, (e.g., text, video, audio, multimedia) and computer-mediated communication (CMC) systems (e.g., e-mails, Internet Relay Chats, MOOs, I-phone, NetMeeting) which create new and rich teaching/learning environments (Pastor, 1998; Ryan et al., 2000).

From this perspective, network technology is used for the delivery of instruction (Martin, 2005). This means that the educational process is experienced online, in different virtual learning environments (VLEs). In this paper, the World Wide Web is considered a culturally constructed environment (Baym, 1995; Lawley, 1994; Reid, 1995). Under this concept, it is the space where social relations occur as well as the tool that people use to enter this space. Furthermore, it is the space where communities are formed, when individuals meet and share common beliefs and practices not in physical reality (as in face-to-face communication) but by means of communication technology, in a virtual world (Jones, 1995; Jordan, 1999). Thus, beyond an educational tool, Internet provides a cultural space, the cyberspace, where people can interact, share common knowledge and information and, build communities in different virtual environments.

VLEs are currently appearing for the purposes of sign language learning and teaching. Many websites present different sign languages through digital video recordings (streaming video) or through animated signing avatars for vocabulary, grammar, syntax and Deaf culture learning. In this paper we attempt to delineate the process of how a VLE can be formed for the purposes of sign language learning. Particularly, we introduce a VLE for studying British Sign Language (BSL) at the Centre for Deaf Studies, in Bristol University, U.K. The paper will also highlight the factors that influence the design of a VLE as well as the importance for adopting a learner-centred approach when teachers are faced with the challenge of implementing VLEs for sign language learning.

2. CYBERSPACE AND SIGN LANGUAGE LEARNING

2.1 DEFINING VIRTUAL LEARNING

By definition, virtual (or distance) learning is the learning that takes place when (Whiteman, 2002: 4):

1. Teachers and learners engage in learning separated by distance.
2. The instruction is delivered via websites and CMC.
3. The communication is interactive while teachers and students send and receive their feedback.
4. The delivery of the instruction can take place simultaneously or asynchronously.

Virtual learning supplements traditional classrooms with web-based material and learning activities, discussion boards on different course sites and online chats in asynchronous or synchronous way. Thus, the structure of a VLE resembles the one of a traditional environment: "There [is] an interaction space like a classroom, where the "teacher" or others may "lecture" and where group discussions may take place; a communication structure like "office hours", where student and teacher may communicate privately" (Hiltz, 1994: 6). In a conventional classroom, educators and students come together in a physical room for discussion, lectures and tests, whereas in a virtual classroom individuals "come together as telepresences for instruction" (Tiffin & Rajasingham, 1995: 10).

In traditional settings, other resources in addition to the teacher are important (e.g., textbooks, white boards and other laboratory equipment), whereas in a VLE different kinds of media make available information on one web page (Ryan et al., 2000). This feature makes VLEs flexible tools as they enable tutors to support different teaching and learning styles (Waring & Boardman, 2004).

2.2. VLE FOR SIGN LANGUAGE LEARNING

The majority of Deaf/sign language related sites incorporate digital video (video streaming) and animated signing avatars (Verlinden et al., 2002) for delivering information in sign language. Since research has determined that Deaf people have poor reading and writing skills on spoken languages, information on the web is supported by visual material on simple and friendly interfaces (Debevc & Peljhan, 2004). Printed information (e.g., closed captions) consists a L2 to Deaf people and poses many obstacles on understanding the information presented online (Lee et al., 2004; Ohene-Djan et al., 2003), whereas online digital video helps Deaf people in understanding the information, while the material is presented in the language they are familiar with (Debevc & Peljhan, 2004).

In line with this, recent research (Lee et al., 2004) addressed the issue of creating Signing Webs; that is a collection of Web pages created by a particular signing community (e.g., American Sign Language, British Sign Language) and linked together by sign language based connections. The idea is to create Web pages based on video signing where singing links allows hyperlinking within video and/or animation. The development of such pages is based on the incorporation of an authoring tool, the SignEd, in order to author pages in sign language.

By considering the above definition of VLE, a virtual environment for sign language teaching and learning is the combination of Web pages with sign language content and CMC (e.g., e-mails, video conferencing) for online communication and collaboration (Figure 1). Deaf and hearing people meet online to a virtual space, like a website, in order to gain and share knowledge about a target sign language. There, the instruction is delivered mainly by video clips and text as well as through CMC.

Usually, the video clips offer lectures in sign language surrounded by supplemental text-based content (Lee et al., 2004). Some videos have subtitles, which students with good sign language skills can follow and learn the language they do not understand (Debevc & Peljhan, 2004). Online video window (and animation window) has basic video control buttons (play, stop, pause and rewind), so that teachers and students are able to watch the video by jumping to specific scenes (backwards and forwards), which are interesting or difficult to understand.

A VLE is formed at the Centre for Deaf Studies (CDS), in Bristol University, U.K. for BSL learning purposes. In particular, both tutors and students (Deaf and hearing) are provided with access to the following two main web resources for BSL learning, which are both developed in the CDS:

- SignStation: www.signstation.org
- DeafStation: www.deafstation.org

Although students and teachers work mainly with these two, there are other BSL websites, which they retrieve for their learning.

SignStation is a website devoted to people who want to know more about BSL and Deaf people in the workplace. It offers a complete BSL course and it consists of an online BSL dictionary with sign-search facilities; an interactive BSL course "The Company", which includes dialogues, vocabulary, grammar explanations and interactive exercises (Figures 2-3); interactive sign awareness video "A-Z of sign"; and a test yourself quiz with multiple choice questions.

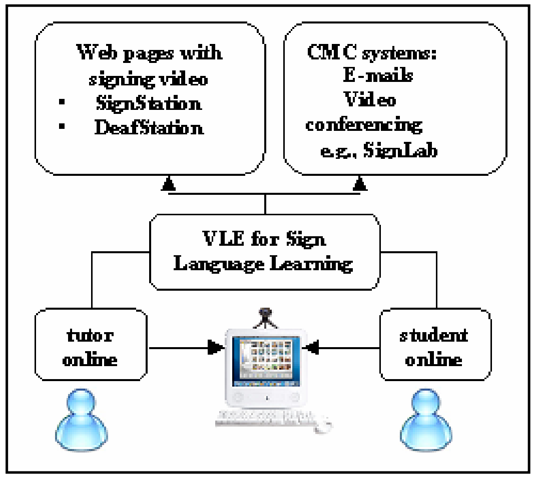


Figure 1: VLE overview for sign language learning

"The Company" is a teach -yourself BSL course based on the theme of a Deaf Design and Building company. Using a storyline of a young Deaf person joining the company and working through to a full scale, on-site building project, the videos show dialogues of increasing complexity. With each dialogue there are illustrated grammar explanations, actual exercises for students to carry out as well as further information on Deaf people and sign language (Figure 2). Like all language courses, there is also a vocabulary list which students can consult at any time and which they can compare with what is being signed in the dialogues.

One of the key features of SignStation is the possibility to search for a sign through a full dictionary of BSL, which can be accessed by text, or through a unique picture interface for different categories. By clicking on the objects in the picture, students can reach the corresponding sign (Figure 3). There are over 5,000 signs available online, which users can use from a mobile phone too.

DeafStation is a website where teachers and students have access to authentic Deaf produced video material concerning news, sports, health, travel, entertainment, and humour. This is a large zone of Deaf material that is updated almost everyday. Especially the news operates everyday, Monday to Friday.

These websites can be used either during classroom time or off-site, when it is convenient to students. Based on current practices, tutors use these sites to supplement parts of their courses and assign to their students project work, where students are required to identify and find certain information about Deaf culture and BSL (e.g., BSL poetry video clips) or analyse BSL in terms of grammar and syntax. In this case, students watch over and over the video clips and then gloss their content.

VLE in the CDS is also formed by the use of a videoconferencing system, SignLab, which is based on the network connection of Apple Mac computers to a central 360 GB server within CDS. In essence, SignLab is a collection of folders, the Home Directories, the Staff Homes, the Teaching Resources and the Staff Private. These folders are on the server and everyone identifies his/her workplace when logging in SignLab.

This system works with Panda software, which is developed exclusively at the CDS and it is used for asynchronous videoconferencing between students and tutors (Figure 4). This is a fundamental element of BSL courses with powerful social dimensions within the department and the University of Bristol. Panda enables recording and posting video messages to each other, in sign language, in an asynchronous way.

Panda-produced materials are of two kinds: (a) video clips prepared by the teachers by filming themselves signing or by digitising old VHS videos (e.g., conversations between Deaf individuals, BSL stories); (b) conversations between tutors and students relating to specific tasks. In this case, students film themselves signing and post their work to their tutors. Then, tutors assess students' work by inserting their comments into the students' original video clips. All Panda-video material is stored into the Teaching Resources folder in a common server, where students have access and retrieve them according to tutors' indications.

This VLE offers a set of basic and complementary services, which are integrated into the teaching and learning of BSL:

- Direct learning with Web materials; students are able to access useful information for BSL learning, which helps to organise and structure the content of their learning without dealing with problems of disorientation or 'information overload' (Pastor, 1998).
- Interactivity for self-learning and assessment; Students have the opportunity to directly interact with their particular activities as well as with their tutors on specific activities through Panda. The latter is the most important feature of VLE since it enables question - answer exchanges between tutors and students on particular tasks. In addition, students can search and observe for their own answers from resources available (e.g., Web and Panda resources) and thus, construct BSL knowledge.
- Student and tutor technical support. There is already evident the need for more ICT training for tutors and students before using this VLE. Therefore, at the beginning of each academic year students get trained to use Panda and SignLab for their courses. Tutors are also supported by technical staff when encounter problems with developing and creating material with Panda and other technology.

3. BIBLIOGRAPHICAL REFERENCES

Baym, N.K. (1995). The Emergence of Community in Computer - Mediated Communication. In S.G. Jones (ed.), *Cybersociety. Computer-Mediated Communication and Community,* (pp. 138-163). California: Sage Publications.

Debevc, M. & Peljhan, Z. (2004). The Role of Video Technology in On-Line Lectures for the Deaf. *Disability and Rehabilitation*, 2; 26(17), 1048--59.

Hiltz, S.R. (1994). *The Virtual Classroom. Learning Without Limits via Computer Networks.* Norwood, New Jersey: Ablex Publishing Corporation.

Jones, S.G. (1995). Understanding Community in the Information Age. In S.G. Jones (ed.), *Cybersociety. Computer-Mediated Communication and Community,* (pp. 10-35). California: Sage Publications.

Jordan, T. (1999). *Cyberpower. The Culture and Politics of Cyberspace and the Internet*. London, New York: Routledge.

Lawley, E.L. (1994). The Sociology of Culture in Computer-Mediated Communication: An Initial Exploration. http://www.itcs.com/elawley/bourdieu.html

Lee, D.G., Richards, J., Hardman, J., Soudain, S., & Fels, D. (2004). Creating Sign Language Web Pages. In: J. Klaus, K. Miesenberger, D. Burger, & W.L. Zagler (Eds), *Computers Helping People with Special Needs. 9th International Conference, ICCHP 2004, Paris, France, July 7-9, 2004, Proceedings.* (pp. 1088-1095) Berlin: Springer.

Martin, M. (2005). Seeing is Believing: The Role of Videoconferencing in Distance Learning, *British Journal of Educational Technology*, 36(3), 397--405.

Ohene-Djan, J., Zimmer, R., Gorle, M., & Naqvi, S. (2003). A personilisable Electronic Book for Video-based Sign Language Education. *Educational Technology and Society,* 6(4), 86--99.

Pastor, E. (1998). Web tools and Web applications. In F. Verdeho & G. Davies (Eds.), *The Virtual Campus. Trends for Higher Education and Training* (pp. 207-210). London: Chapman and Hall.

Reid, E. (1995). Virtual Worlds: Culture and Imagination. In S.G. Jones (ed.), *Cybersociety. Computer-Mediated Communication and Community.* (pp. 164-183). California: Sage Publications.

Ryan, S., Scott, B., Freeman, H., & Patel, D. (2000). *The Virtual University. The Internet and Resource-Based Learning.* London-Sterling (USA): Kogan Page.

Tiffin, J. & Rajasingham, L. (1995). *In Search of the Virtual Class.* London and New York: Routledge.

Verlinden, M., Tijsseling, C., & Frowein, H. (2002). A Signing Avatar on the WWW. In I. Wachsmuth & T. Sowa (Eds.), *Gesture and Sign Language in Human-Computer Interaction: International Gesture Workshop, GW 2001, London, UK, April 18-20, 2001. Revised Papers* (pp. 169-172). Springer – Verlag Berling Heidelberg. http://www.springerlink.com/media/7panqnxxxq49a1t79trn/contributions/0/f/b/d/0fbdcv691m2qcw9d.pdf

Waring, M., & Boardman, K. (2004). Learning to Teach, Teaching to Learn: A Developmental Framework for Teacher Training. In D.S. Preston (ed.), *Virtual Learning and Higher Education* (pp. 1-13). Amsterdam-New York: Rodopi B.V.

Whiteman, JA.M. (2002). The Adult Learner in the Computer Mediated Environment. *ERIC Document Reproduction Service (EDRS)*, ED 467889, CE 08372

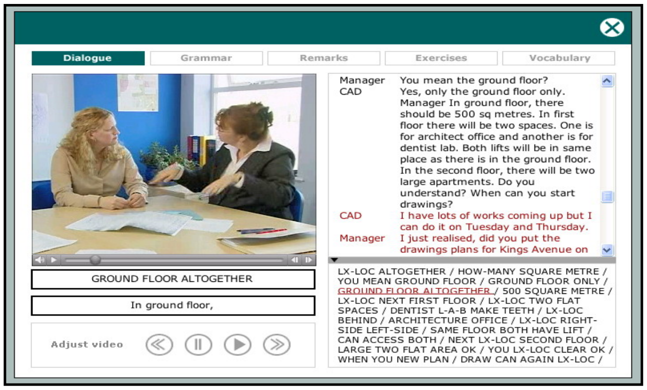


Figure 2: Sample of BSL course online

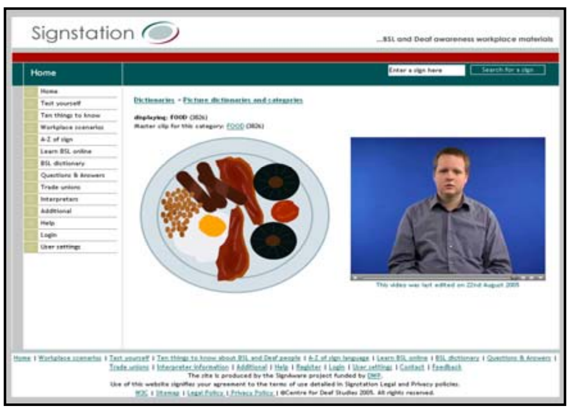


Figure 3: Sample of BSL dictionary online

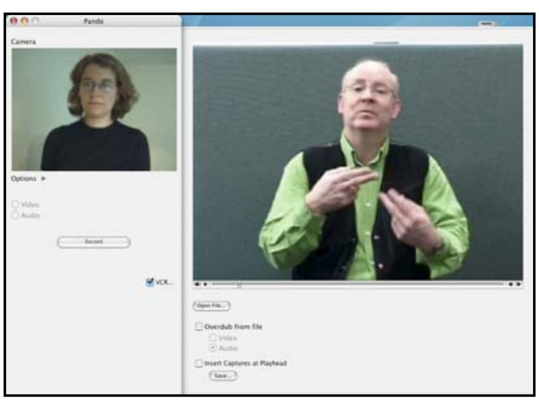


Figure 4: Panda window with video file loaded

PILOT LEARNING STRATEGIES IN STEP WITH NEW TECHNOLOGIES: LIS AND ITALIAN IN A BILINGUAL MULTIMEDIA CONTEXT 'TELL ME A DICTIONARY'

E. Insolera, M.G. Militano, E. Radutzky, A. Rossini

LIS.Me.Di.A.
Via D. Bramante, 43 - 05100 Terni
info@mpdf.it

Abstract

A pilot project designed for the integrated or non-integrated classroom, speech therapy setting, and family at home, this multi-media DVD + book series offers deaf and hearing children "of all ages" a lively interactive tool for discovering and comparing two very different languages, Italian Sign Language and Italian.

"Raccontami un dizionario"(Tell Me A Dictionary) is rich in vocabulary presented through stories and sentences that project both languages as living languages, thanks also to a lively 8-minute animated cartoon, signed and spoken narration, Italian with subtitles, vocabulary building games and a glossary that takes you back to the vocabulary items in the DVD.

The illustrations and story in the accompanying book derive from the DVD: both animated and printed versions tell the story even without the support of language, permitting access even to young children just beginning to read.

The animated story facilitates the understanding of written Italian, especially verbs which, through animation, offer a dynamism that is limited by two dimensional book illustrations. The book reinforces the written Italian and children can experiment narrating the story to their friends and engage in dramatization with classmates.

Published by LisMedia & CO, it is easily adapted to other spoken and signed languages.

1. PROLOGUE

Making the learning of a new language enjoyable is not always easy to accomplish. It is even less so if you have been deprived the possibility of acquiring a mother tongue, that language that is assimilated naturally if the child is exposed to it at an early age.

All too often, deaf children travel down a scholastic and speech therapy path in which the only language encountered is vocal language, like spoken Italian or spoken Greek.

Vocal language is not natural for deaf children, since it travels on the acoustic-vocal channel that is inaccessible to them because of their deafness; for this reason the path is arduous, and the objective is reached only after hard work, enormous effort and frustrations for the deaf child.

Even written language, though visual, is no small challenge.

A highly effective way to favour integration that aims at equality through access, participation and respect is to permit deaf children to measure themselves with the hearing children, taking advantage of their natural language, sign language, as a strong point and not merely as a marker of diversity.

The desire to see sign language find an appropriate place within the teaching and speech therapy contexts with the need for true integration of deaf and hearing children, coupled with the need for a specific tool that favours learning LIS and Italian in a bilingual context, gave rise to a pilot interactive multimedia series, "Raccontami un Dizionario" (Tell Me a Dictionary") and its first volume "*Un Picnic Tutto Pazzo*" (A Crazy Crazy Picnic) and accompanying book.

This series offers deaf and hearing children an interactive and enjoyable instrument that permits them to discover and compare two very different languages, and offers them a stimulating venue to "grow" together.

At a first glance, this material appears to be designed for very young children, because of its clarity and its lively primary colours throughout.

As one uses the material, one realizes that each section is multi-layered, so there are semantic and grammatical features to suit everyone's age and reading abilities, or lack of.

It is currently being used by kindergarten children, elementary school and high school students, parents, and teachers of sign language courses for hearing and deaf non-signing adults.

It is being used in classrooms where there are no deaf children, but where hearing children are learning LIS as a second language.

It has been acquired by parents whose hearing children are learning sign language because, for one reason or another, they are unable to communicate through speech.

At first glance, this material might seem difficult for the younger children, who are just beginning to read, since the language presented in the DVD is at times complex.

But, thanks to the wonderful world of multi media technology, this series seems to turn upside down the classic learning readiness rules for certain aspects of language.

Children grow with the material, and like Alice in Wonderland, the level of understanding will increase with the age of the child.

It is an expensive product, and so it had to be designed to be used over and over, and for a long period.

This pilot DVD and accompanying book were realized by LIS.Me.Di.A. & Co., a consortium composed of the Istituto Statale di Istruzione Specializzata per Sordi, the Mason Perkins Deafness Fund, and Digisys, s.r.l., with a contribution from the Comune di Roma (Municipality of Rome).

2. CHARACTERISTICS

"Un picnic tutto pazzo" has several characteristics that, together, render it somewhat unique:

- The animated cartoon facilitates the comprehension of the accompanying written subtitles, in particular, the verbs that, by way of animation, conserve their dynamic state. It is extremely difficult to understand new verbs when reading books, especially because verbs are difficult to draw, because there are never enough illustrations to match the text, and because the illustrations are static and two-dimensional. Even more difficult to understand in children's books in the absence of movement are relationships of cause and effect;
- The illustrations and the images in the DVD are extremely clear and colourful, and they enable the child to comprehend the story even before and without reading the text and without knowing how to read;
- The two languages, LIS and Italian, are presented through **single words**, that are placed in different types of **sentences** (negative, interrogative, affirmative, imperative, etc.) that contain the words, the story and the detailed account filled with interesting and humorous detail - all aimed at enriching the lexical, syntactic and narrative aspects of both languages;
- The signs and words are placed in a context (e.g. picnic) and subdivided into categories (e.g. food, beverages) to favour rapid vocabulary learning and a memorization that will last over time;
- The story is quite humorous and action-packed, aimed at richer language and enticing the children to return to the DVD over and over; the sentences in the dictionary section too, where possible, offer not only a context but are humorous and educational;
- Taking advantage of the fact that the letters of the computer keyboard appear in upper case, but when the child pressed on a capital letter, a small case equivalent appears on the screen (if the cap lock is not on, of course). We have chosen to use upper case in the more elementary dictionary section and the lower case in the subtitles that narrate the story in written Italian. The book, too, uses lower case written Italian.
- The DVD is technically designed to be seen interactively either on the computer screen, or on a large television screen with the whole class or the family at home, using the remote control.
- With the games section of the DVD, children can test what they have learned and recover what they have missed;
- The small book that accompanies the DVD is designed to root the written Italian once the story has been completely understood. The book version of the story offers children the possibility of reflecting, on what they are reading. And since it is a small book, they can take it anywhere, even on the bus, and they can also experiment the pleasure of narrating the story to their friends and relatives.

3. DVD CONTENT AND METHOD

3.1. Racconto Animato (The animated story)

This animated cartoon is about a 'crazy, crazy' experience that Lino and Milli, brother and sister, have in their futile attempt to have a picnic. The story is told first through animation (with the option of a voiceover and subtitles). In order for the story to be understood by all children, even very young children who have not yet learned how to read, the characters communicate with very clear actions and expressions and no words or subtitles are necessary for total comprehension.

Since comprehension of the story facilitates learning the signed and written narration, the cartoon section should be viewed one or more times before visiting the others.

The very first time the children view the animated cartoon; it is most effective to deactivate the volume to allow children to fully concentrate on understanding the cartoon.

The volume can be activated during subsequent viewings to let those children who have access to the spoken language practice the written Italian reinforced by the voiceover.

3.2. Racconto LIS (Story in LIS)

The Racconto LIS should be the second section to be viewed.

Added to what was experienced in the first section, here the "simpatico" deaf actor Emilio narrates the story of the cartoon in Italian Sign Language (LIS), right alongside the cartoon, which is now viewed in a television screen in what is supposed to be Lino and Milli's bedroom.

Emilio narrates the entire story line, whereas in the following section, he will reappear, offering a detailed description of what the children are seeing in addition to the story line.

3.3. Racconto LIS Descrittivo (Detailed Description in LIS)

In this descriptive section, Racconto LIS Descrittivo, the entire cartoon has been divided in small segments, and indexed so the child or facilitator can pick the segment she wants to concentrate on.

Each segment is shown just as pure animation, with no voiceover and no Italian subtitles.

Then the segment that the child has just viewed stops and Emilio appears on the screen to recount the segment in great detail, adding rich vocabulary and grammar, a myriad of classifiers, all easy to understand after having viewed the animation.

Children can click on the segment again and again or pass on to another segment that they wish to view. Emilio's detailed description in LIS can also be viewed with Italian subtitles and heard through voiceover by activating the appropriate button.

3.4. Disegni Segni Parole (Drawings Signs Words)

We have designed a descriptive dictionary section named Disegni Segni Parole that is rich in drawings, signs and words, where you can learn signs, words, and then you can see the signs and words in sentences.

The interactive "dictionary" section is comprised of six categories, each containing 48 everyday key vocabulary words, all concrete objects that the children will have already encountered in the animated story.

The child first selects an object. Each object is represented by a written word in Italian, but also by the drawing of the object so that children who do not yet know how to read can still choose the object without understanding the meaning of the word.

When a child clicks on the drawing of an object, the object pops out and then Emilio appears on the screen to show how to sign it in LIS. The same word is then shown and heard in written and spoken Italian respectively.

One can also access the translation in English of the Italian word equivalent.

Clicking on Emilio, he will repeat the sign, and the child can keep clicking until he has learned the sign.

Clicking on the link "*frase*" ("sentence"), the same sign appears, this time in a complete sentence that is in turn translated into both written and spoken Italian.

Just before the sentence appears, there is animation on the screen that helps the child who does not read to understand the meaning of the sentence to come.

The sentence, like the single sign, can be seen over and over.

This is an enjoyable while remarkably instructive way to provide children with sentences and vocabulary within a context.

Care has been taken to provide sentences that are at times humorous, at times educational, and that provide a wide variety of grammatical usage.

It goes without saying that those important classifiers abound, since the narration for the most part derives from the animated cartoon.

It is already evident to us that this section of the DVD must be amplified to include hundreds of key vocabulary words and sentences in the subsequent volumes.

3.5. Giochiamo (Let's Play)

This game section, Giochiamo, was designed to enable the users of the material to test their comprehension of single signs.

Emilio signs an object and children have to choose the correct object or Italian word from a list of words and objects displayed at Emilio's side.

Milli informs you if your answer is correct, and if not, Lino, being 'psychologically correct', encourages you to try again.

To prevent clever kids from guessing correctly because they have memorized the place the words occupy on the list, we keep mixing the order.

A colourful board of apples keeps the score and whenever you answer correctly, another apple is bitten. (Which, by the way, this is the section that is visited most frequently by adults!)

3.6. Glossario (Glossary)

In this section, Glossario, the letters of the ABC alphabet are presented with the equivalent handshapes from the LIS manual alphabet. You can click on any letter to access a large number of words that are used throughout the narration and in the sentences from the dictionary section.

The vocabulary items listed are strictly tied to a specific context and cannot necessarily be generalized.

This reflects one of the peculiar features of sign languages, obviously because they are visual-gestural languages. When you click on a word, the word is shown within a sentence, thus clarifying it's meaning and illustrating the context in which it can be used.

Further context is offered in parentheses alongside many of the vocabulary items that require disambiguation.

For example, when reference was made in a sentence to a mother and her baby ducks that swim, it was necessary, next to the word for 'swim' in the glossary a parenthesis indicating that it was for duck swimming, and in a particular way in which it appeared in the sentence context. (This is clearly delicate, for we are dealing with glosses, which no researcher likes. We realize that we can never be careful enough. In the future we must elaborate a way to include the very useful glossary in which there is no possible way that the user might apply the vocabulary learned to contexts in which that particular sign is not acceptable.)

Each sentence is shown twice to promote comprehension.

The vocabulary items that appear in the Glossario are presented in three colours:
yellow for the words/signs that are the 48 key words, green for those appearing in the key word sentences but are not the 48 key words, and **blue** for those appearing within the cartoon.

All of the vocabulary words appear in bold and are appropriately emphasized in the subtitled sentences, respecting the proper colour for each of the three types.

3.7. The Book that accompanies the DVD

The use of the book is strongly recommended only after children have watched the DVD one or more times.

Once the children have totally understood the story in the DVD, they are ready to learn the written Italian version in the book, and discuss it or narrate it to others with greater ease.

Both the DVD and the book are tailored to "children of all ages". Since it is humorous and so rich in grammar, it is now very popular even with adult sign language classes.

4. TESTING THE MATERIAL

The material was tested by several teachers in kindergarten and elementary school and by bilingual speech therapists.

Reports were very positive, which was expected since even the smallest pool of water found in the desert is lifesaving!

In the testing phase, teachers were delighted to see how the children were riveted to the screen.

Most importantly, they were impressed by how especially pleased and participatory the deaf children in the viewing group were to see their language on the TV or computer screen, in a primary role.

Teachers also commented that the children delighted in the interactivity and in being able to learn on their own, without adult assistance, promoting independence of choice.

They recounted that the story of Lino and Milli and the group of vindictive ants whose vacation has been ruined by the two children in search of a good place to have their picnic attracts the children to the point where they are able to protract their interest and willingly engage in activities that the teacher creates from the DVD.

The DVD motivates them through a clear stimulus at the visual and linguistic levels, and represents an indispensable basis for any teaching and learning activities.

It involves the student in a repetition of the lexicon and sentence structure in various forms: play, narration, exercises and favour a remarkable “natural” acquisition.

Teachers relate that they are all too aware of the effort on the part of their deaf students to memorize new words and with what difficulty they are able to understand in which contexts to use them and how these materials contribute to reduce the effort.

We are told that this DVD is easy to consult, the images and signs are an effective support in memorizing Italian words, and enrich the lexicon of their classes, while thoroughly enjoying it, in LIS and in Italian.

They remarked at how easy it was to work on the sign language and the Italian once the story was totally comprehended through animation.

The most gratifying feedback, perhaps, was learning that teachers used the DVD to spin off various art, language and dramatization activities, which we hope in the future to collect and publish as a teachers’ resource anthology.

Feedback from the actual users and facilitators is crucial to continual improvement as subsequent volumes in the series are created.

5. BATTLE OF THE LANGUAGES

Creating this pilot was truly stimulating because we had no real precedents and our task was to bring together two languages that travel on different channels, and combine them with the languages of the multi media that would, in the end, transmit the content of the DVD to the users.

One simply cannot fathom the amount of time (and near fistfights) involved in arriving at comprehension among the different “languages” and media.

The next attempt, also with the help of feedback and further testing, will surely be easier.

The most important thing we learned was not to begin with the storyline in animation, but in LIS, and design the animation after the LIS narration, avoiding the need to narrate, often too fast, to fit the scenes.

At times, when the resulting LIS production for a scene was unacceptably too rapid, we redid the animation for those scenes, but this proved much too time-consuming and expensive.

We are now brainstorming a storyline for the next DVD. Any suggestions are welcome, especially from deaf storytellers!

Metodo VISTA – Teaching Sign Language in Italy

Emanuela Cameracanna, Maria Luisa Franchi
Gruppo S.I.L.I.S. and Mason Perkins Deafness Fund MPDF
Via Nomentana, 54 – 00136 Rome – Italy
ml.franchi@fastwebnet.it – emyariete@libero.it

Abstract

The **Metodo VISTA** is a video course consisting of a Teacher's Book and a Teacher's Video, a Student's Book and a Student's Video. It is based on the book **"Signing Naturally"** for the teaching of American Sign Language written by **Chery Smith, Ella Mae Lentz and Ken Mikos**, and has been adapted for the teaching of LIS. Thus the first, second and third volume are intended for teachers who wish to teach LIS and for students who want to learn it. Its aim is to help teachers organize a series of lessons divided into three different levels of language learning. The **Metodo VISTA** leads the students who know nothing about deafness or Sign Language to interact with deaf people in a wide range of situations. The knowledge of the culture of the deaf is an integral part of the programme. It is also taught by the presentation of native signers who show cultural and linguistic behaviour in various situations in a video.

The S.I.L.S. Group and the Mason Perkins Deafness Fund (MPDF) have collaborated since 1997 and published:
Cameracanna, E. Franchi, M.L. Rossini, P. (1997). Metodo VISTA 1° livello. Roma (It): Kappa publisher.
Cameracanna, E. Franchi, M.L. (2000). Metodo VISTA 2° livello. Roma (It): Kappa publisher.
Cameracanna, E. Franchi, M.L. (2003). Metodo VISTA 1° livello. Roma (It): Kappa publisher.
The **Metodo VISTA** is a video course consisting of a Teacher's Book and a Teacher's Video, a Student's Book and a Student's Video. It is based on the book **"Signing Naturally"** for the teaching of American Sign Language written by **Chery Smith, Ella Mae Lentz and Ken Mikos**, and has been adapted for the teaching of LIS. Thus, the first, the second and the third volume are intended for teachers who wish to teach LIS and for students who want to learn it. Its aim is to help teachers organize a series of lessons divided into three different levels of language learning.
Each level is structured in five or six units, each of which is divided into chapters and comprises the organization, activities and material of the unit. Furthermore, each Teacher's Book contains one unit dedicated to language review, supplemented by a video on cultural behaviour, values and social rules. The **Metodo VISTA** leads the students who know nothing about deafness or Sign Language to interact with deaf people in a wide range of situations. The knowledge of the culture of the deaf is an integral part of the programme. It is also taught by the presentation of native signers who show cultural and linguistic behaviour in various situations in a video. Before devising the programme, various approaches to foreign language teaching were analysed in search of an approach which stresses interpersonal communication. In the end, an approach was chosen that focuses on the communicative intentions during daily interaction among people, such as, for instance, introducing oneself, inviting somebody to one's house, asking somebody a favour, discussing a film, a book or an event, telling a tale or an adventure.

1. THE PHILOSOPHY OF TEACHING

The units refer to the following principles:

1. Students learn a language better when the contents of the units are presented in a context;
2. Students memorize the language more easily when the activities are meaningful and based on personal experiences;
3. Students develop comprehension skills faster than production skills.

Therefore, all units are devised to be presented in LIS, avoiding the use of the voice, of written Italian, of "translations" or comments and the teachers are encouraged to sign on a level of production which is slightly superior to the students' level.

2. ORGANIZATION OF THE UNITS

The "list of contents" at the beginning of each unit shows the teacher at a quick glance what is to be treated in each lesson.
Each unit is divided in basic sections:

GYMNASTICS:

It supplies the teacher with useful activities to help the students warm up their hands and to teach them the proper use of their hands in order to sign correctly.

PERCEPTION AND VISUAL DISCRIMINATION:

The teacher is provided with useful activities to accustom his students to "listen" with their eyes, to memorize and to learn to distinguish the signs.

INTRODUCTION:

It supplies useful activities to the teacher in order to work exclusively on the students' comprehension, concentrating on the way of presenting the grammar and new words of a specific topic in a context.

SIGN PRODUCTION:

It shows the teacher how to train the students in developing the production of signs. The production activities focus on sentence structure, dialogues or narrative practice.

INTERACTION:

It supplies the teacher with communicative activities ranging from structured to spontaneous dialogues, allowing the students to integrate grammatical structures they have studied with new words. The students become acquainted with an interlocutor's reaction and experiment with conversational strategies.

EXTENDED COMPREHENSION

It shows the teacher how to support and extend the topics, grammar and vocabulary presented in the chapter. This activity is useful to increase comprehension skills. Stories told by the teacher should be a regular part of the lesson. The narration of experiences or events which happened in class will lead the students to develop comprehension strategies without necessarily understanding every sign.

COMPREHENSION CHECKS:

They allow the teachers to check the comprehension of concepts, topics or linguistic structures which were treated in the foregoing lessons. Obviously Level 2 provides for a larger number of activities of this kind than the precedent level.

NARRATIVE PRACTICE:

Provides teachers with comunicative activities that help develop tecniques for narrating. Students develop skills for role shifting, spatial structuring, sequencing events and establishing time references.

BREAKAWAYS:

They provide the teacher with the possibility of varying the activities, to develop relationships in class and to work on other aspects of LIS. They are divided into six types of activities: group activities, facial expressions, visualization, giving orders, dactylology and numbers and mime. These activities can be used by the teacher at random any time and can be repeated as needed.

TEAM-BUILDING:

Activities of this kind foster relationships and lower the anxiety level in class

FACIAL EXPRESSIONS:

These activities provide the students with some more practice to improve both the grammatical and emotional non-manual components

VISUALIZATION:

These activities help the students to develop their skills in the use of space and glance

GIVING ORDERS:

These activities provide the students with yet another way of memorizing words, grammar and functions

FINGERSPELLING AND NUMBERS:

These activities train the students to master these specific aspects of the language

MIME:

These activities train the students in the development of miming capacities which help them to master the non-manual components of the language

3. TEACHER PREPARATION

We suggest at the teacher to many counsel , for exemple th following:
View the teacher's video which is designed to demostrate various activity types and discuss language principles central to the curriculum.
Read the entire curriculum to familiarize yourself with how functions, topic and grammar features are sequenced and recycled throughout the level.
Familiarize yourself with the Transcription Conventions to help you understand the glosses in the text should be memorized or a prepared in avance so that you do not have to read and sign simultaneously in class.

4. TRANSCRIPTION CONVENTIONS: SYMBOLS USED FOR WRITING LIS

Symbol	Example	Explanation
WORD	DEAF	An Italian word in capital letters stands for a LIS sign (this word is called a gloss) The meaning of the sign and the Italian word may not be exactly the same.
#	#CAR	A pound sign (#) indicates a fingerspelling loan sign

Etc.

CLARIFICATION OF SOME PUZZLING GLOSSES

IX	short for INDEX, IX indicates **pointing** and is used for third person pronouns (he, she, it, him, her) Specific referents are indicated by italicized words in quotation marks, immediately following the gloss.
IX-dir	**IX-**dir is used when the pointing gives **directions** or traces a route to a place (i.e., IX-dir "*around the corner to the rigth*").

Etc.

SYMBOLS FOR NON MANUAL BEHAVIOR

Symbol	**Example**	**Explanation**
------q	-----------q	
	YOU EAT FINISH	yes/no questions
Etc.		

SYMBOLS FOR CLASSIFIERS

Symbol	**Explanation**
DCL	descriptive classifier sign used to describe an object or a person. What is described is italicized and quotation marks (i.e. DCL*"curly hair"*). Sometimes referred to as size and shape specifiers or SASSes.
Etc.	

ORGANIZATION OF LEVEL 1

Each of the six units focuses on different basic subject as follows:
Unit 1: How to introduce yourself
Unit 2: Exchange of personal information
Unit 3: Description of the surroundings
Unit 4: Talk about where you live
Unit 5: Talk about your family
Unit 6: Everyday life – what we do and when
Cumulative review: units 13-17

Each unit uses and integrates the topics of the subject, the vocabulary and grammar presented in the precedent lessons. The students learn to exchange personal information, such as give their name and address and talk about their families.

ORGANIZATION OF LEVEL 2

Each of the six units focuses on different basic subject as follows:
Unit 7: Giving directions
Unit 8: Describing others
Unit 9: Making requests
Unit 10: Talking about family and occupations
Unit 11: Attributing qualities to others
Unit 12: talking about routines
Cumulative review: units 7-12
Each unit uses and integrates the topics of the subject, the vocabulary and grammar presented in the precedent lessons. The students learn to exchange personal information, such as give their name and address and talk about their families.

ORGANIZATION OF LEVEL 3

Each of the six units focuses on different basic subject as follows:

Unit 13: Complaining, making suggestions and requests
Unit 14: Describing and identifying things
Unit 15: Locating things around the house
Unit 16: Talking about the weekend
Unit 17: Exchenging personal information: life events
Cumulative review: units 13-17

ORGANIZATION OF CUMULATIVE REVIEWS

This unit is organized such as to repeat all dialogues of the foregoing lessons and is based on the presentation of dialogues in the video.
The video shows cultural behaviour as well as topics which are important for relationships regarding exchange and relations among different cultures and also conversational strategies. The students learn to introduce themselves properly, how to draw attention and how to participate at a sign conversation and watch deaf people telling stories, which helps them to develop role switching skills, which, though difficult to learn, are most important in narration.

VIDEOTEXTS AND WORKBOOK LEVEL I, II AND III

Are conceived to revise and put into practice what has been learned in class as well as to increase the comprehension of signed narrative. Each student should have a copy of the *Videotexts and workbook* these materials provide students with a way to review, prectice and retain what they have learned in class. Our Experience with Sign Language books led us to conclude that video is a most effective medium for sign language materials. We found that students use books as a reference for remembering signs, but unfortunately, most Sign Language books associate Italian with signs. This defeats the purpose of the curriculum, which encourages students to think only in sign.
There are many advantages in using a video and workbook:

- In the video you can see the movement of the signs exactly as they are used in the sentence.
- You can see the way in which the shape of a sign is influenced by the preceding or following sign
- You can learn how the movement of a sign can be modified to change the meaning
- You can see how a facial expression appears
- You can observe how the movement of the body, of the head and the eyes can be used to open or close a sentence and to express grammatical concordance
- You can see how language is used in a context
- You can see how a visual language like LIS can be used creatively in poetry, theatre and narration
- Last but not least you can watch the video as often as you wish
- The video are devoid of audio.

Each unit concentrates on one of the main functions of language, such as introducing oneself, talking about one's family or everyday life – what we do and when we do certain things, how to complain or make requests,

describe objects, speak about furniture, talk about the weekend, about events of life or family life.
By means of this functional approach the language learned is the one that is used in everyday conversations. Learning the functions of language in an interactive context also leads to the improvement of conversational skills by using various aspects of time (recurring and continuous), inflected verbs and role-switching, as well as various kinds of classifiers and references to space.
The units are divided into various sections:

CONVERSATIONAL PRACTICE

Each unit of the video begins with some conversations which, in the workbook, are accompanied by dialogues that stress the functions of language and key expressions.

GRAMMAR NOTES

The notes on grammar explain the grammatical structures introduced in the unit and may also include some notes on the narrative structure, the principles referring to sequence and to various shifts in a story. They are usually followed by the subsequent section in which they are applied.

COMPREHENSION EXERCISES

These exercises are comprehension activities to test the mastery of the vocabulary and linguistic functions you have learned.
Other activities are based on answering questions, making summaries or filling in blanks in the Workbook. Last but not least pair-work activities are a means of putting into practice what you have learned together with a partner.

GRAMMAR EXERCISES

The parts of the video containing grammar exercises draw our attention to grammatical structures in a pre-established context, demonstrating specific aspects of grammar.

CULTURAL/LINGUISTIC NOTES

In order to understand a language you have to understand its cultural context. The cultural/linguistic notes provide a view on the history of values and social rules of the community of deaf people.

STORY CORNER

This is a story, either humorous or simply informative, created and filmed for the purpose of training reception skills, learning vocabulary in a context as well as developing strategies to grasp the meaning of the stories even without understanding all signs.

CONCLUSION

We hope that by using this book, in the future students of LIS will be able to sign naturally, to interact in a relaxed way and that they will get to know and respect the community of deaf people. The aim of this manual is to preserve the integrity of the language and to encourage more deaf people to become teachers of LIS. This programme, together with linguistic research on sign language in process and original works of literature produced within the community of deaf people, will continue to demonstrate that LIS is an ingenious, elegant and effective expression of a fascinating culture.