

Handling SL annotations of the handshapes

typannot, Dominique Boutet^{1,2}, Claudia S. Bianchini³, Claire Danet⁴,
Patrick Doan^{4,5}, Morgane Rébulard⁴, Adrien Contesse⁴, Léa Chèvrefils-Desbiolles²

1 : Dynamique du Langage In Situ (DyLIS) *Université de Rouen Normandie* 7, rue Thomas Becket 76830 MONT SAINT AIGNAN Cedex - France

2 : Université de Rouen Normandie (URN) *Université de Rouen Normandie* 1, rue Thomas Becket 76821 Mont-Saint-Aignan Cedex - France

3 : Formes et Représentations en Linguistique et Littérature (FoReLL-EA3816) *Université de Poitiers Maison des Sciences de l'Homme et la Société* 99 avenue du Recteur-Pineau 86000 Poitiers - France

4 : Amiens School of Art and Design (ESAD Amiens) *Ministère de la Culture et de la Communication* 40, rue des Teinturiers 80080 Amiens - France

5 : Université de Technologie de Compiègne [Compiègne] (UTC) *Université de Technologie de Compiègne [Compiègne]* rue du Dr Schweitzer Compiègne, 60200 FRANCE +33 3 44 23 44 23 - France

Objectives of the transcription system

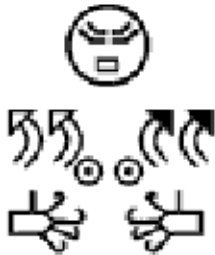
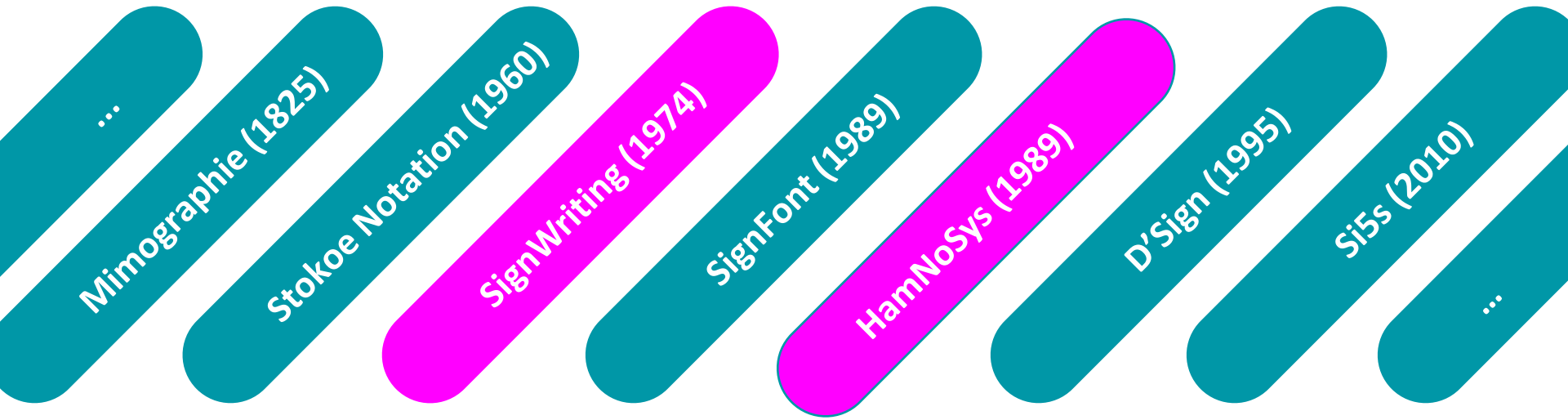
Elan - HSG1.eaf

File Edit Annotation Tier Type Search View Options Window Help

The screenshot displays the Elan transcription software interface. At the top, there is a menu bar with options: File, Edit, Annotation, Tier Type, Search, View, Options, Window, and Help. Below the menu bar, there are two video windows showing a conversation between two men. The left window shows both men, while the right window is a close-up of one man gesturing. Below the video windows, there is a timeline and a selection bar. The timeline shows a time range from 00:00:16.000 to 00:00:29.000. The selection bar indicates a selection from 00:00:17.370 to 00:00:17.470. Below the selection bar, there are playback controls and checkboxes for Selection Mode and Loop Mode. The main part of the interface is a transcription grid. The grid has columns for time and rows for different tiers of transcription. The tiers are: RH ID gloss, LH ID gloss, Dependent variabl, 1-handed/2-handed, Grammatical func, HS preceding targ, HS following targ, Genre of text which, and Summary. The grid shows various annotations and their corresponding time intervals. For example, the LH ID gloss tier shows annotations like 'SAME', 'SCOTLAND', 'HOLD-FLAG', 'PTloc', 'G', 'PEOPLE-LOOK', 'NORTH', 'IRELAND', 'SAME', 'PEOPLE-LOOK', and 'PTpro1'. The grid also shows annotations for other tiers, such as 'HO PTto G', 'PE N IRELA SA PEOP', 'PT BEF P', 'PT YOUTH', 'c', '1', '2', 'd', 'o', 'p', '5', '1', 'c', 'c1do', 'X1', 'c2', '51', and '5'.

Nr	Annotation	Begin Time	End Time	Duration
1	SAME	00:00:15.140	00:00:15.330	00:00:00.190
2	SCOTLAND	00:00:15.330	00:00:15.630	00:00:00.300
3	HOLD-FLAG	00:00:17.040	00:00:17.270	00:00:00.230
4	PTloc	00:00:17.270	00:00:17.590	00:00:00.320
5	G	00:00:17.590	00:00:17.890	00:00:00.300
6	PEOPLE-LOOK	00:00:19.380	00:00:19.640	00:00:00.260
7	NORTH	00:00:19.650	00:00:19.860	00:00:00.210
8	IRELAND	00:00:19.870	00:00:20.310	00:00:00.440
9	SAME	00:00:20.320	00:00:20.550	00:00:00.230
10	PEOPLE-LOOK	00:00:20.550	00:00:20.930	00:00:00.380
11	PTpro1	00:00:24.420	00:00:24.610	00:00:00.190

Representing Sign Languages









Limits of existing transcription systems

Notation systems must meet 5 main criteria to make possible accurate analyses and their reading :

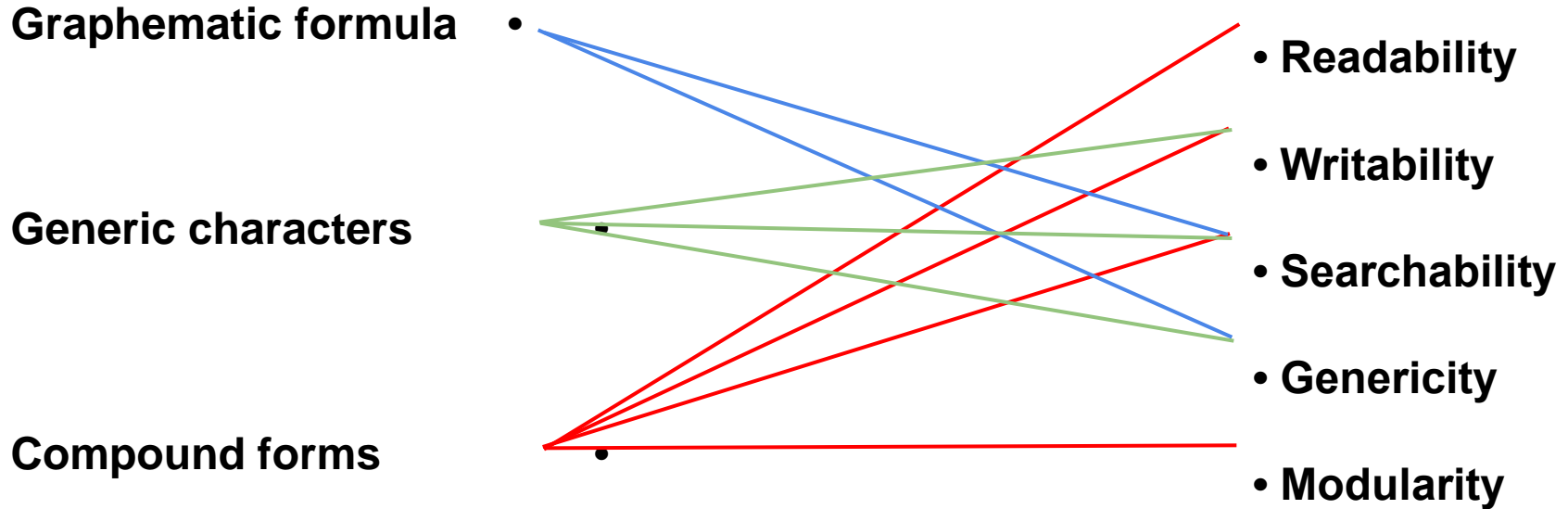
	Readability		Writability		Searchability	Genericity	Modularity
	Human	Computer	Human	Computer			

Limits of existing transcription systems

Notation systems must meet 5 main criteria to make possible accurate analyses and their reading :

	Readability		Writability		Searchability	Genericity	Modularity
	Human	Computer	Human	Computer			
SignWriting	<input type="checkbox"/>	 <input type="checkbox"/>	<input type="checkbox"/>	 <input type="checkbox"/>	 <input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
HamNoSys	 <input type="checkbox"/>	<input type="checkbox"/>	 <input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	 <input type="checkbox"/>

TYPANNOT uses 3 levels of representation



Level 1: the graphematic formula



1TOCscCO 234Fo

1	Index
TO	Thumb Opposed
Csc	Curved semi-closed
CO	COntact
234Fo	Middle Ring Pinky Flat open

Level 2: the generic characters

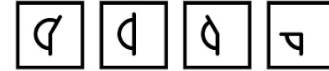
Lateralization



Finger designation

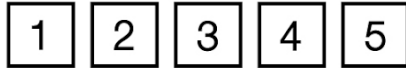


Angle



Finger selection

Counting



Clustering



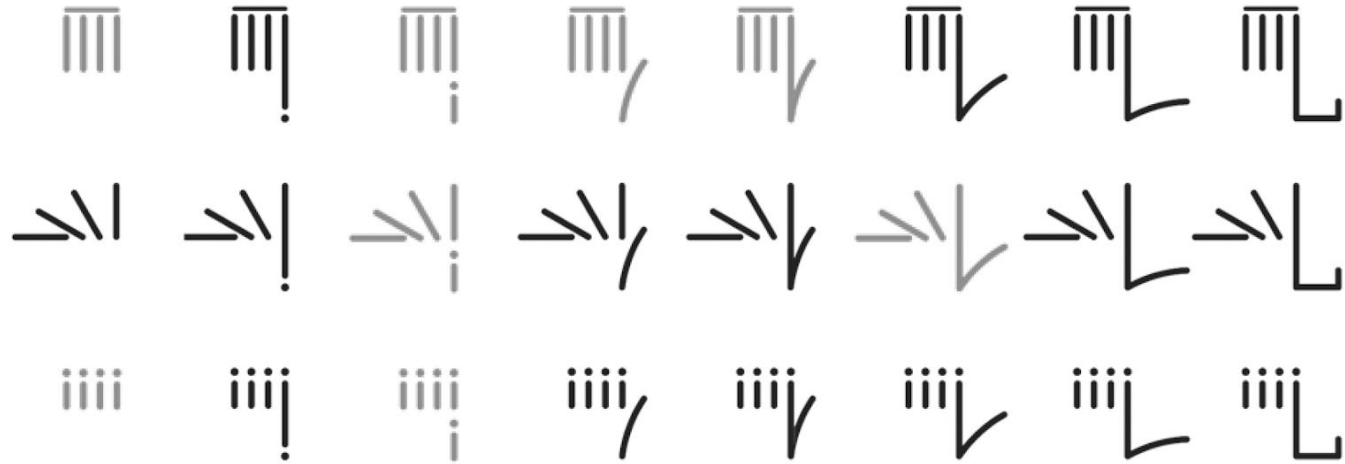
Closeness



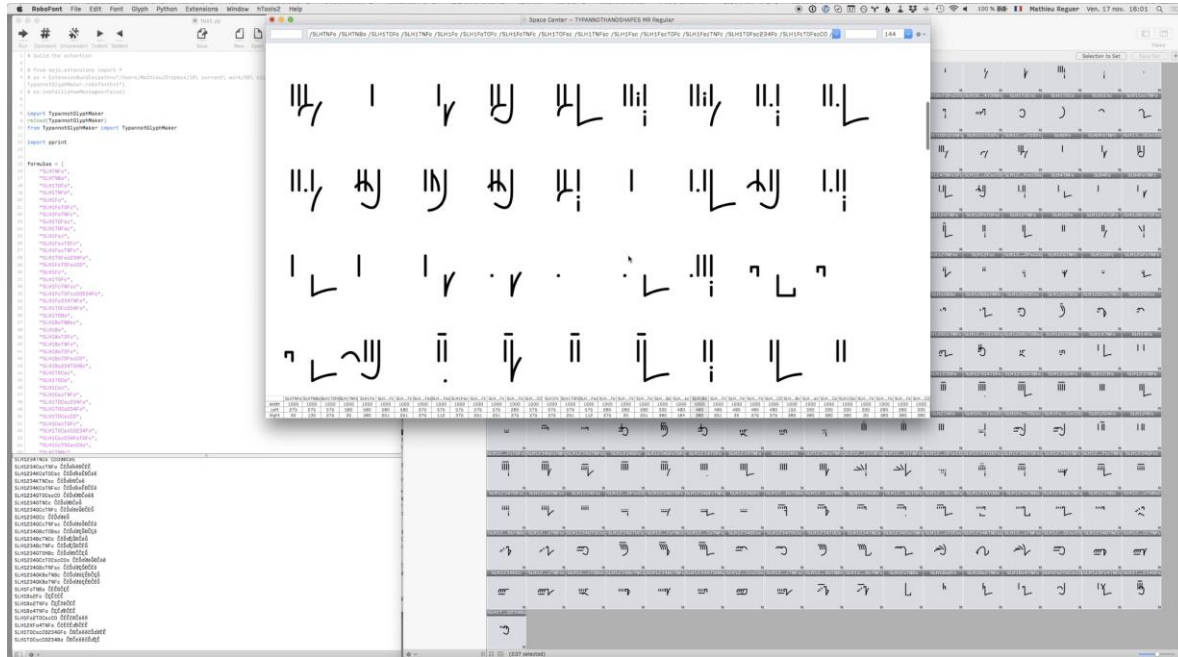
Shape



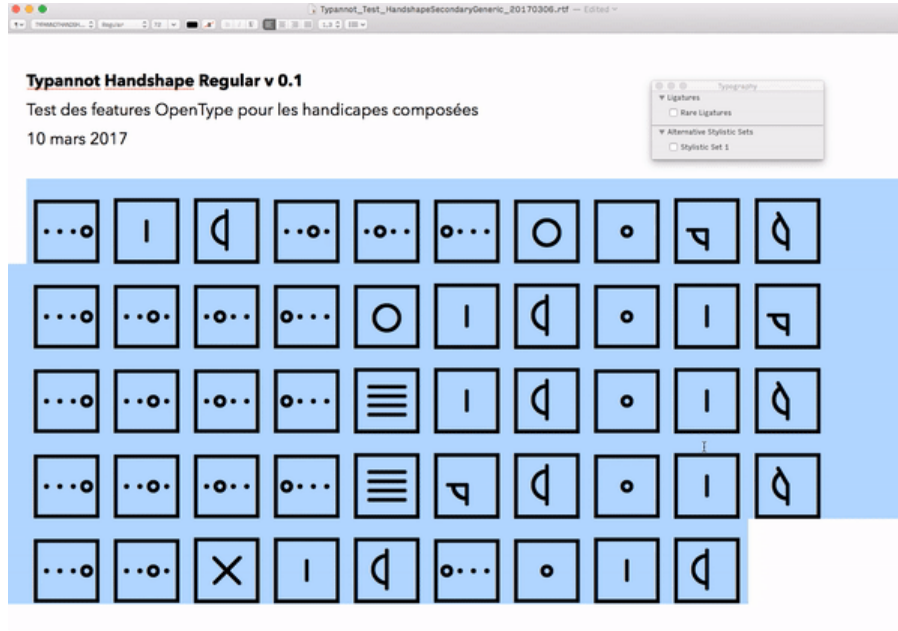
Level 3: the compound forms



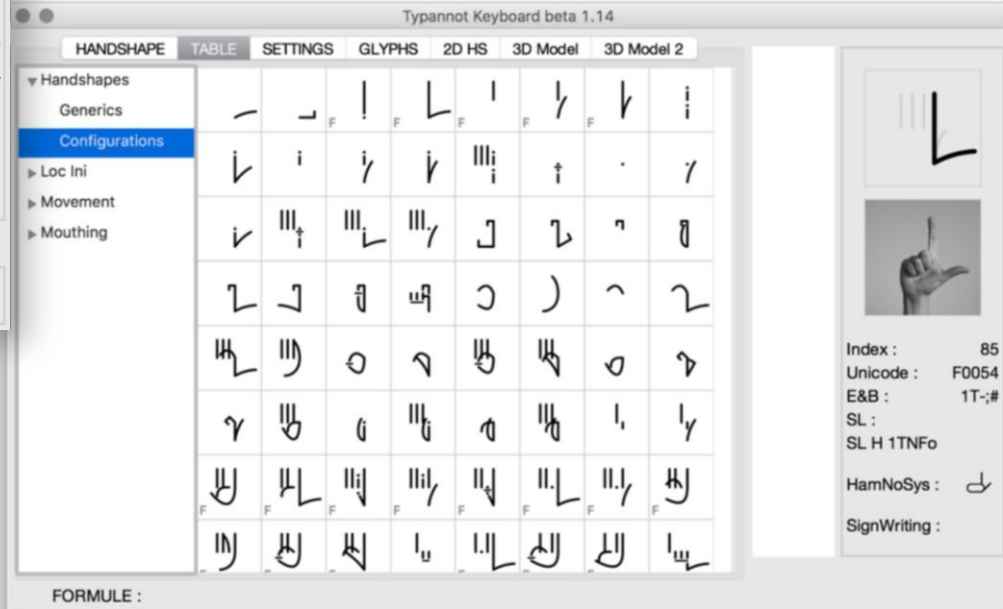
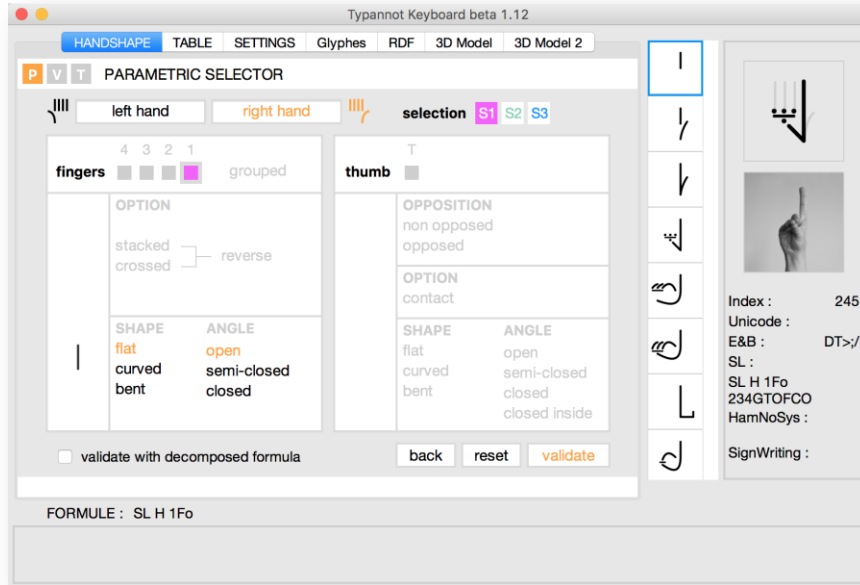
Compound forms are generated



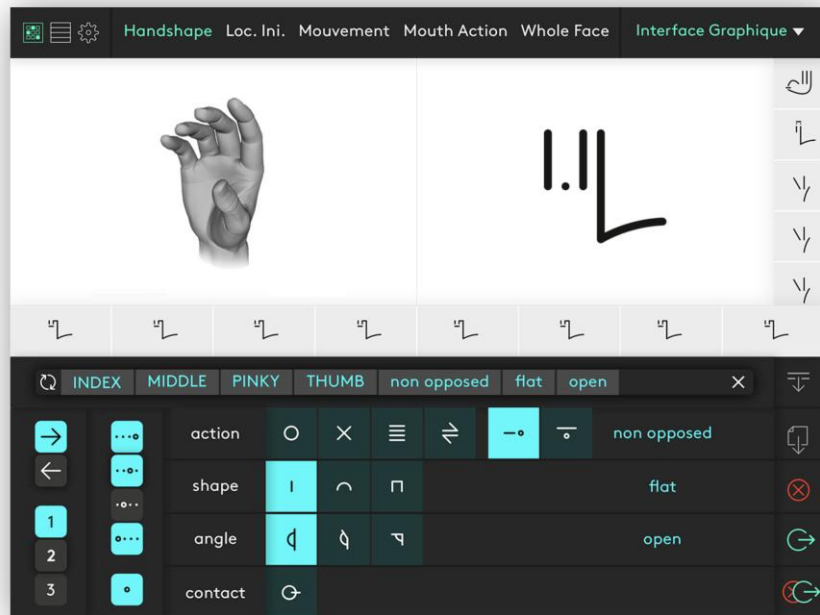
Two ways to look at it



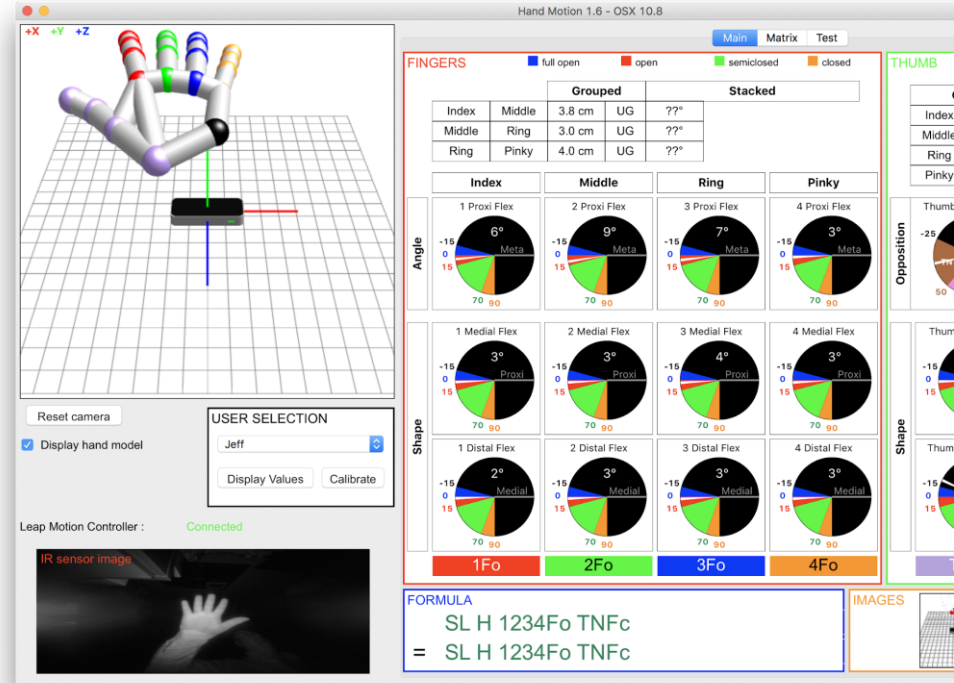
Input interfaces



Work in progress

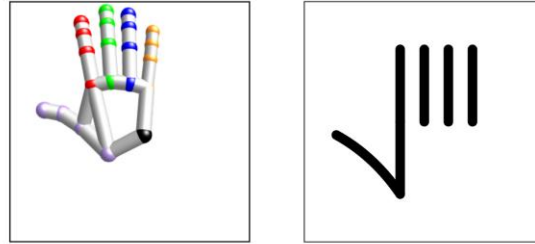


Gestural Interfaces



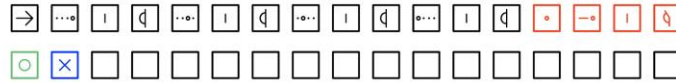
Gestural Interfaces

Typannot Signs 1.0



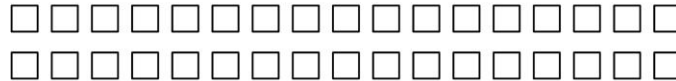
R1Fo92Fo93Fo94Fo9TNFscG99999X99999K99999COT9c

PRIMARY GENERICS

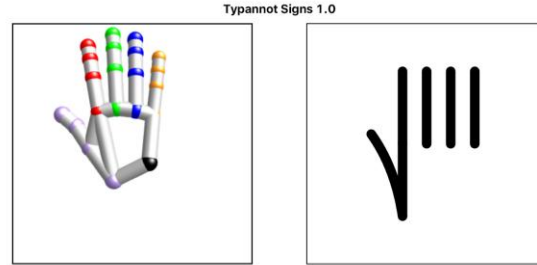


1234Fo TNFsc

SECONDARY GENERICS

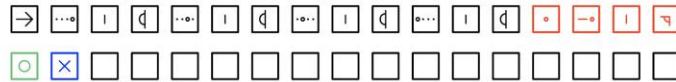


Gestural Interfaces



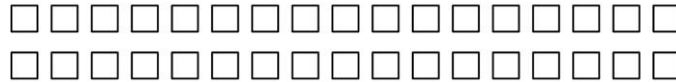
R1Fo92Fo93Fo94Fo9TNFc9G99999X99999K99999COT99

PRIMARY GENERICS

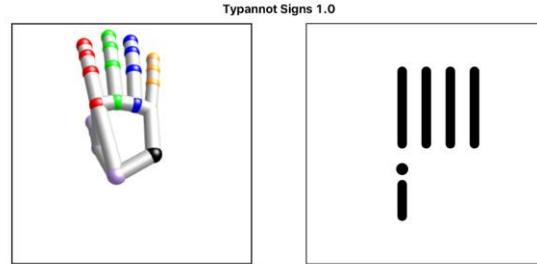


1234Fo TNFc

SECONDARY GENERICS

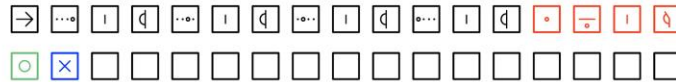


Gestural Interfaces



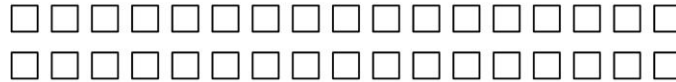
R1Fo92Fo93Fo94Fo9TOFscG99999X99999K99999COT99

PRIMARY GENERICS

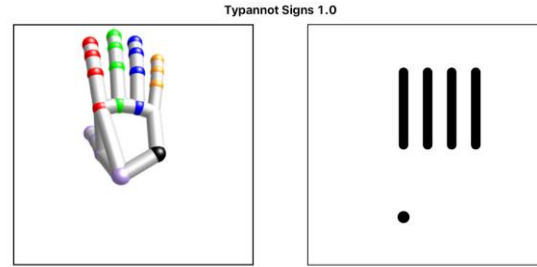


1234Fo TOFsc

SECONDARY GENERICS

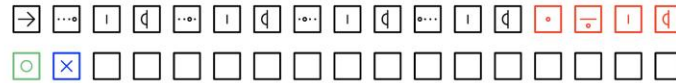


Gestural Interfaces



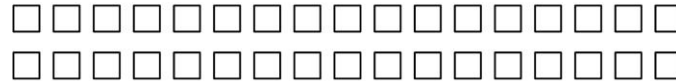
R1Fo92Fo93Fo94Fo9TOFo9G99999X99999K99999COT99

PRIMARY GENERICS



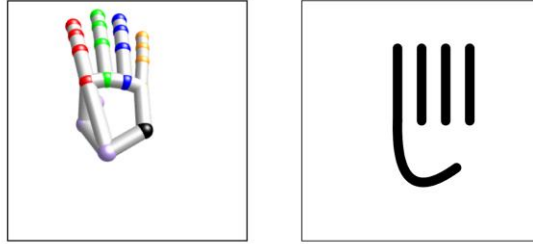
1234Fo TOFo

SECONDARY GENERICS



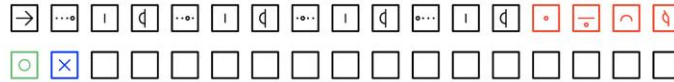
Gestural Interfaces

Typannot Signs 1.0



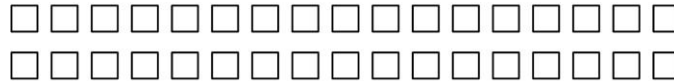
R1Fo92Fo93Fo94Fo9TOCscG99999X99999K99999COT99

PRIMARY GENERICS



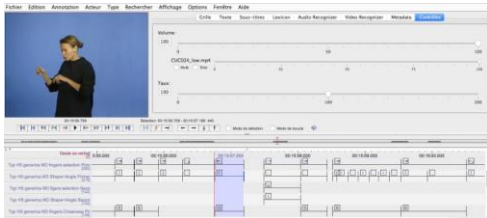
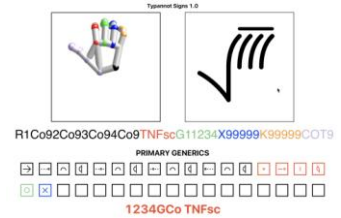
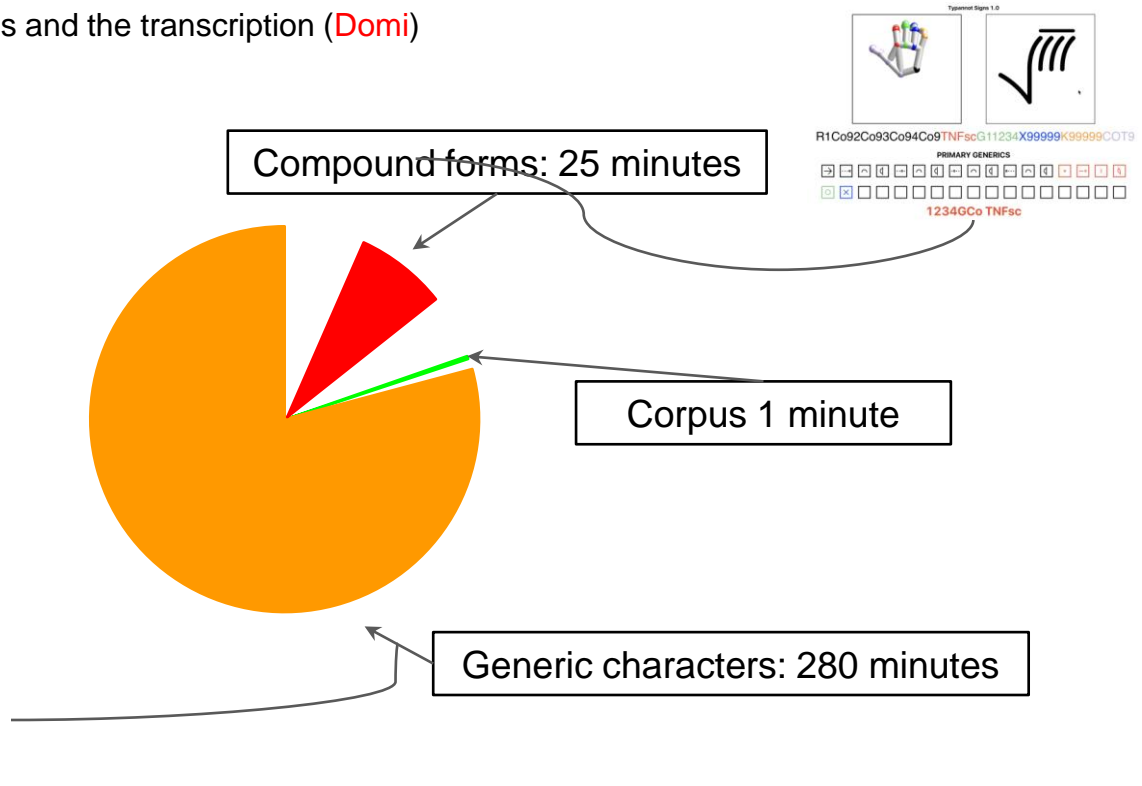
1234Fo TOCsc

SECONDARY GENERICS

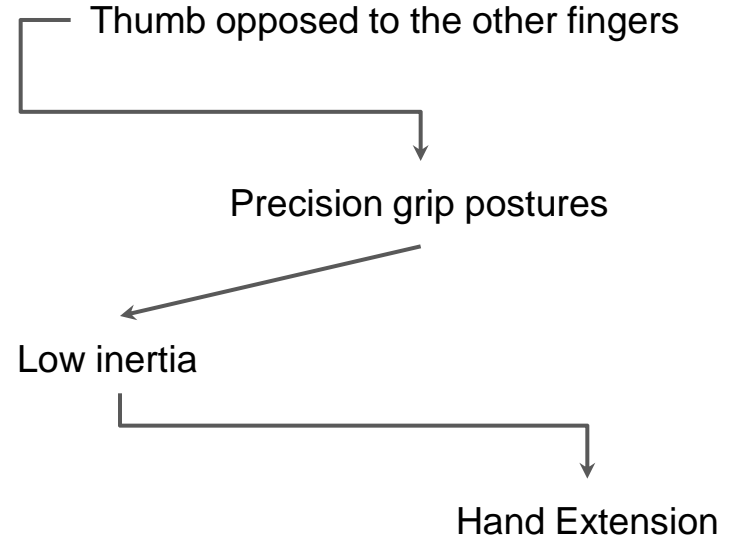
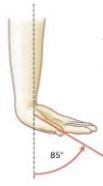
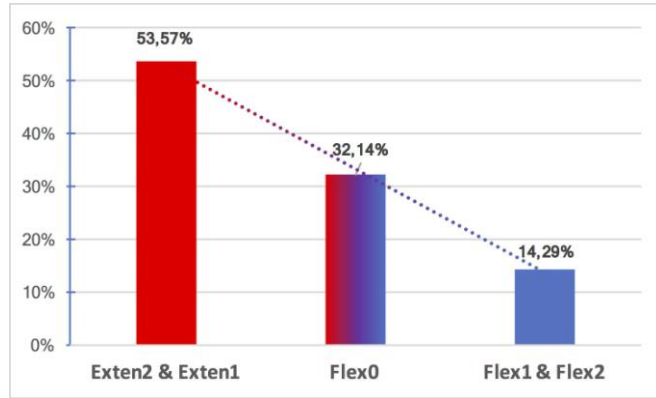


Transcription time in Elan

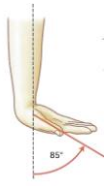
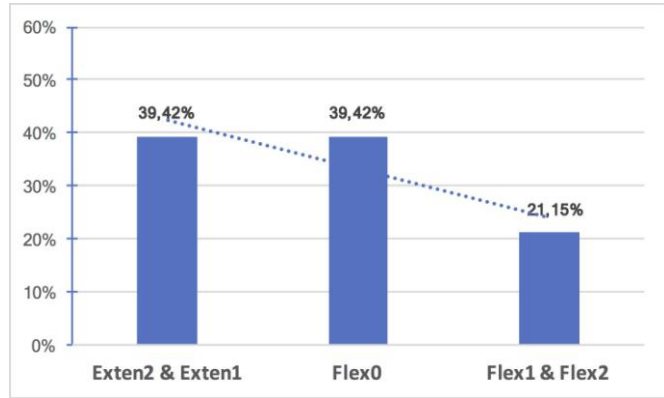
Some clues of the ratio between a corpus and the transcription (**Domi**)



Some results: Praxic gestures/symbolic gestures



Some results: Praxic gestures/symbolic gestures



Thumb **NON** opposed to the other fingers

Power grip postures

Strong inertia

Less Hand Extension, more intermediate locations



Perspectives

HANDSHAPES (2018)

- The gestural interface,
- The virtual keyboard (graphical and keyboard interfaces)
- Transcription of ½ hour of LSColin (5000 signs)

LOC INI (2019)

- A set of the generic characters
- A set of compound forms
- A minimal gestural interface
- Transcription of ½ hour of LSColin

MOUTH ACTION (2019)

- A set of graphematic formula
- A set of the generic characters
- A set of compound forms
- Transcription of 10 minutes of LSColin

01
02

Studying SL requires to transcribe SL video recordings into a written form, here using Elan, to deconstruct language into searchable data. There's a difference between transcribing and translating:

- A transcription is when you change the modality from oral to writing
- A translation is when you go from a one language to another

As we can see in this picture, when researchers want to transcribe SL, what they actually do is that they first interpret it and then they transcribe it into their own language using glosses. And the usage of glosses comes down to translating a sign into a glosse. To put it simply, It is as if to analyse the structure of a Chinese speech, you were studying an English written translation. Linguists have to do that because there's not of a sufficiently effective writing system to transcribe SL.

And what SL Researchers really need, is to stay as close as possible to the original signal to be able to understand the underlying structure of those languages.

To do that, they need a dedicated writing system able to break up Sign Languages into their smallest structural components.

And to be efficient, this writing system need to have a typographic system that fits specific criteria. To explain the difference between writing system and typographic system, the writing system defines the value of each glyph, and the typographic system defines the shape of each glyph.

03

During this presentation we are going to talk about Sign Languages. Sign languages are the visual and gestural languages used by deaf people. The meaning of those languages is given by the hands, the face and the whole body of signers.

Most of the languages that exist in the world don't have their own writing system. Many languages can share a similar writing system and each for each of them, some small customizations can be made to fit the language. But the same thing isn't true with sign languages. The structure of those languages is so different from other oral languages that they can't efficiently use alphabets or any other writing system created for spoken languages.

Since the 19th century, many attempts have been made to create a notation system for sign languages. Amongst them, we can highlight SignWriting and Hamnosys.

Long story short, SignWriting is an education oriented system. The different components of language are written using iconic glyphs.

Hamnosys on the other hand is a system used by linguists to transcribe Sign language corpora. It's partially iconic and mainly focuses on hands components.

05

If we look at the two existing system that we talked about we can see that SignWriting is user friendly but doesn't work well with computers... HamNosys does, but it is far from being user friendly.

An other thing we can say: none of them fits very well the genericity and modularity criteria.

What we are going to show you today is our work on a new transcription system called Typannot.

It is developed by a team of linguists, type and digital designers and developers. This transdisciplinarity is the foundation of our work method which enables us to target the highest levels of fulfilment for each of the 5 criteria.

We are building a system that will make transcription easier and shorter for linguists --- a system by which all signs components (handshapes, placement, movement, and facial actions) can be transcribed --- and a system that will enable the searchability of glyphs and all there individual linguistic components.

Focusing on our work on handshapes for today's presentation we are going to show you how Typannot works.

06

Typannot is based on three levels of representation:

- A graphematic level
- A generic level
- A composed level

Each of those three levels offer specific answers to our 5 criteria. We are going to see how.

07

First : The Graphematic formula is an organised list of all components needed to describe handshapes. Each of those components is identified by an alphanumeric code. It makes it easy to type on a keyboard and it gives total searchability.

The structure of that list as been worked on an improved within time in order to get the most accurate description of each component --- and to be fully efficient computerized data --- which ensures genericity.

08

Then: The generic characters is nothing else than an iconic representation of graphematic codes. Those icones are much easier to learn, recognise and use (and I'm talking here for human beings).

One icone for one graphematic code in order to keep the data organisation optimized.

09

Finally: The compound forms. Based on the graphematic formula these compound forms represents all the components with only one sign.

At the beginning of each line is a handshake, followed by several possible variations.

This typographic system works with modular elements, and those individual typographic elements are direct representations of our graphematic codes.

10

The signs that you see on this picture where generated with our modular typographic pieces and put together by using computerized typographic ligatures.

11

Thanks to this system, transcriptions using Typannot can be looked at on both the generic level and on the typographic level.

Opentype technology allows a simple switch between 2 forms of representation while keeping the raw data available for searchability. And that's pretty handy !

Having done this work on the organisation of data and glyphs was every important. --- But it not enough... because linguists need to tool to use what we have setted up. And that's what we are working on today.

12

We are working a virtual keyboard that will enable linguists to use Typannot in the most easy and efficient way.

Here you can see a prototype by which we have tested the different technological possibilities.

With this keyboard, you can use our generic characters, you can create typographic glyphs in order to send them to ELAN or any other transcription software or you can even go through our glyph libraries.

As useful as it was to set up this prototype, we can't say that the design nore the user experience has been our main priorities until now.

But it is a very important aspect to open our work to the broadest use.

13

That's why we have very recently started working on the design of our future keyboard.

Here is a snapshot of an early work in progress phaze.

When we are working on the design, we are first trying to optimize the placement of each feature, so it feels effortless to find.

We are also working on it's shape, color or type, to ensure it is pleasant to the eye.

And we are working on static and animated feedbacks so that our users will fully understand how the keyboard works and what they can do with it.

We are still a long way from having completed that task, but we're already confident in our capability to create a very efficient and user friendly tool.

Oh and there's a part of our prototype that we haven't talked about yet, and Dominique is going to show it to you.

14

This is one of working gestural interfaces. On the left, you can see the device plugged to the computer. The screen shot on the right is one of our working interface which displays the handshape of the hand. Each column corresponds to one finger. The pie charts display the shape (flat, curved or bent) and the angles (open, semiclosed and closed). We defined some thresholds. It is the case with the thumb for which we distinguish the opposed and the non opposed position, opposed to the other fingers.

You can see the simplified graphematic formula below. And the specific glyph at the bottom right. As I have said, this interface is a working interface.

15

But we are working on several interface to visualize the generics besides specific glyphs. Here you can see the relative position of the thumb changing.

16

and you can see the difference for both specific glyph and generic glyphs.

17

When the thumb is opposed and semiclosed in a flat shape.

18

Or still for a thumb opposed, the thumb still in a flat shape with a more open angle

19

Or when the thumb is curved.

Nowadays the gestural interface is not ready to be used by everybody. But we are working on it.

We all know that transcribe a multimodal corpus is time consuming. The time spend to transcribe is one of the key point of any research.

At the moment for the typannot transcription system, we have three levels of informations :

1/The first one is a graphematic level, it concerns the Loc Ini and the Movement for which we have not any type font (we are working hard to have one)

2/ The second is a generic glyphic system. This level is a type font translation of the features embedded in each handshapes for instance. For now we have generics for handshapes.

3/The third level concerns a final typefont system. We have built just one typefont, this is for the handshapes. Each character of this typefont hold the generics which compose all of the features.

The ratio between the duration of the corpus and the duration to transcribe is very very high if we have to transcribe all of the features for the handshapes for instance.

We have made some tests, and we can say that the ratio is 1/280. It means that it takes 280 minutes to transcribe 1 mn of the corpus. So 1mn of the corpus needs 4hours 40 minutes.

If we use the virtual keyboard with the gestural interface, the ratio falls to 1/25. It means that it takes 25 minutes to transcribe 1 mn of the corpus. In this case, all of the features are included.

To conclue on this point, the modular approach we promote in typannot made by the ligatures of many informations represents large savings in time.

To continue on handshapes, we present some results achieved by the typannot system.

These preliminary results deal with the impact of praxic gestures on symbolic gestures as Signs in a SL. Usually no connection is made between praxic and symbolic gestures. But actually there is no reason to think we can delineate these types of gesturing easily. Deaf people as hearing people use daily gestures to interact with the world. How can we imagine there is no influence from the praxic gestures to the symbolic gestures? Both praxic and symbolic gestures use manual handling. The praxic one undergo objects, weight, inertia, whereas the symbolic handling is fake without exterior inertia and weight. We precisely focus on the consequences of the interaction with the world to determine two types of grasping.

When the thumb is opposed to the other fingers, it corresponds to a precision grip in the praxic gestures. This kind of grip is used to seize or hold a light object. The inertia of these objects are very low. Then the intrinsic location of the hand can be an extension because the object is not heavy. As you can see in the illustration to the right in the slide.

We transcribe the handshapes and the relative location of the hand during 1 minute and 30 seconds of the corpus LSColin in French Sign language. The purpose is to make a proof of concept of the impact of praxic gestures on signs.

When the thumb is opposed in a symbolic gestural system like the corpus we studied, it seems that we see a trend to have an extension location of the hand.

22

When the thumb is NON opposed, like that. It corresponds to a power grip in the praxic gestures as you can see below in the image. This grip is used to seize or hold a heavy object this time. The inertia associated to these objects is very strong. The consequence on the hand is obviously a more balanced distribution of the flexion/extension.

Here again we can see the impact of praxic gestures on symbolic gestures, again with the LS Colin corpus.

The reason is quite clear : we use our gestures to interact with the world everyday. We can expect that this praxic gesture has an impact on symbolic gestures.

We have to investigate further this topic of the influence of praxic gestures on Sign Languages and on cospeech gestures as well. But I think we have the tools to do this.

23

For the Handshapes we would like to achieve the gestural interface during this year and to have a functional virtual keyboard for the handshapes before the end of this year. Then we can transcribe the equivalent of 5000 signs.

During the next year we are going to work on the parameters LOCATION and MOUTH ACTION. The graphematic formula for the LOC INI is already done. We plan to have a set of generic characters in order to transcribe in ELAN. As we have seen with the handshapes, transcribe with the compound forms is much faster. So we would like to develop a set of compound forms, in order to make a transcription of 1/2 hour of LSColin.

Book of Abstracts

Journée d'études AFLiCo JET

Corpora and Representativeness

jeudi 3 et vendredi 4 mai 2018
Bâtiment Max Weber (W)

Keynote speakers

Dawn Knight et Thomas Egan

Plus d'infos :
bit.ly/2FPGxub



Université
Paris Nanterre

Université
Paris Nanterre
UL 220
CIRA
Centre de Recherches Linguistiques

AFLiCo

modys

PARIS8

Dominique Bounet, Claudia Blanchini, Claire Danet, Patrick Doan, Morgane Réboulard, Adrien Contesse, Lea Chéreau-Desbrières (Université de Rouen, Université de Poitiers, ESAD Amiens, UTC Compiègne)

dominique_bounet@orange.fr claudia.blanchini@univ-poitiers.fr claire.danet@gmail.com

adrien.contesse@gmail.com morgane.reboulard@gmail.com patrick.doan@gmail.com leachereau@univ-poitiers.fr

Handling Sign Language annotations of the handshapes

Historically, transcriptions in SL have largely focused on relaying the meaning of a sign rather than the sign's physical shape [1,2]. Two main graphical systems of annotation — SignWriting [3] and HamNoSys [4] — are able to annotate the form of the parameters: the first one is not designed to annotate [5]; the second, in most cases, is coupled to a database tool for integrating SL corpus, named iLex [6]. Therefore, except for few corpora [7,8], the parameters are only partially annotated. One of the reasons is the time needed to do so, even if this has not been documented in literature. Another challenge in the transcription of signs is the integration of generic information (features) into the specifics of each parameter of signs.

At a phonological level [9, 10], the features extracted for handshapes generate multi-million possibilities. Some rules to reduce these potential combinations have been established [11, 12], and others might still be found.

The aim of this communication is to present Typannot [13, 14], a type font for handshapes created to transcribe all the existing SL (142 SL for Ethnologue [15]). This glyphic type font is readable, writable, searchable. It works upon a modular system; the separate features can be assembled into glyphs that

15

enable the representation of every handshape. The features are all searchable through a set of generic (symbolic) glyphs. Those glyphs can be visualized by the annotator at any time thanks to the use of ligatures (OpenType format [16]), and are therefore transparent. At present, this font — freely downloadable on every operating system — can be used with ELAN [17] to transcribe directly by using a virtual keyboard (MacOS) or, for some SL, by using a dedicated template.

Apart from the keyboard and templates for ELAN, a third annotation device has been developed. It works with a motion capture (MoCap) device (based on Leap Motion [18]) by which the annotator can transcribe a corpus by directly using his own hand (real-time controlled). We are in the process of measuring the annotation time per minute of corpus for the different levels and devices with generic glyphs, with specific glyphs and with MoCap. The tendency seems to favour the use of the MoCap device.

Finally, we will present some preliminary results of the transcription made with the Typannot Typefont, at several levels of annotation: from the selection of the fingers, the shapes and the bending angles of the fingers in the palm, their lateral distance to the final glyph of each handshape through the combination shape-angle.