

Ri-organizzare SignWriting
per favorire
la ricerca linguistica
sulle Lingue dei Segni

Re-arranging SignWriting
to promote
linguistic research
on Sign Languages



Claudia S. Bianchini¹, Fabrizio Borgia²⁻³, Maria De Marsico²

1 - Université de Poitiers – EA3816 FoReLL

2 - Università di Roma I «Sapienza» – Dip. Informatica

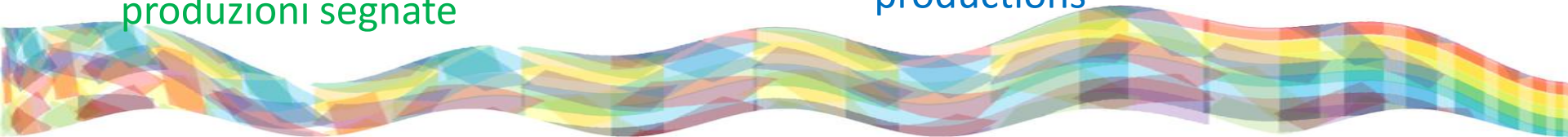
3 - Université de Toulouse III «P. Sabatier» - UMR5505 IRIT

Le Lingue dei Segni (LS)

- Lo status delle LS è assimilabile a quello delle altre lingue esclusivamente «faccia-a-faccia»
- MA... non si può scrivere una LS con l'Alfabeto Fonetico Internazionale
 - a causa della modalità visivo-gestuale
- L'assenza di un sistema di notazione crea difficoltà nella notazione, annotazione e successiva analisi di produzioni segnate

Sign Languages (SLs)

- The status of SLs is comparable to all other “face-to-face only” languages
- BUT... we can't use the International Phonetic Alphabet to write down SLs
 - because of the visuo-gestural modality
- The lack of a notation system hinders notation, annotation, and subsequent analysis of signed productions

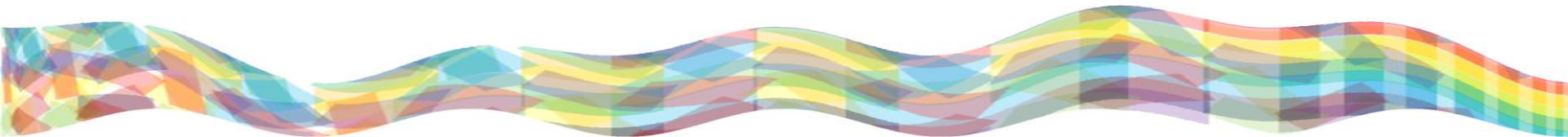


Rappresentare le LS

- Numerose soluzioni sono state trovate per rappresentare le LS
- SignWriting (SW)
 - inventato da V. Sutton ('70)
 - secondo noi, il più valido sistema grafico per le LS

Representing SLs

- Many solutions have been considered to represent SLs
- SignWriting (SW)
 - invented by V. Sutton ('70s)
 - in our opinion, the most valid graphic system for SLs

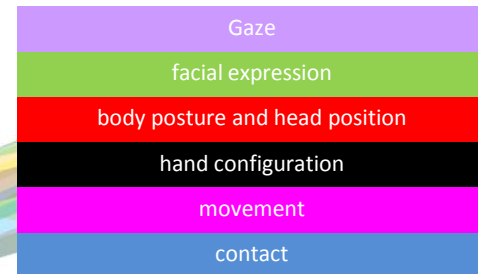


Caratteristiche di SignWriting

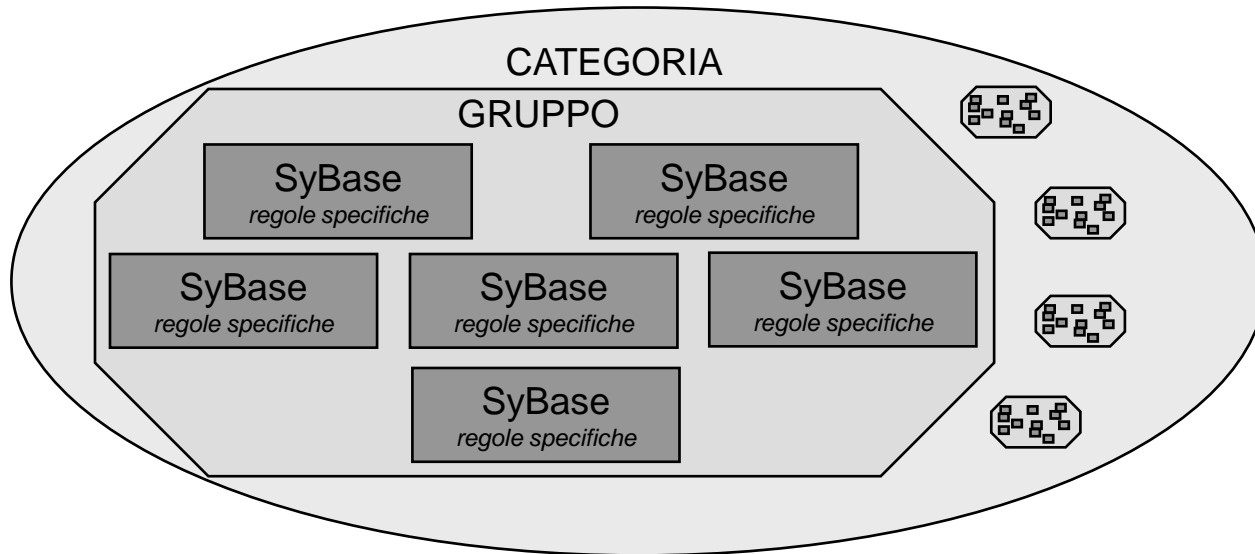
- Rappresentazione analogica 2D dello spazio segnico
- Rappresentazione fedele dei segni
- Possibilità di rappresentare sia le componenti manuali che quelle non manuali dei segni...
- ma... consta di 38.711 glifi diversi

Features of SignWriting

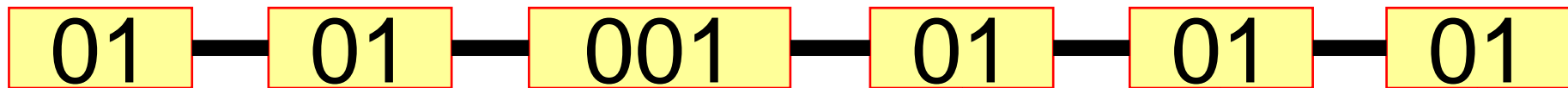
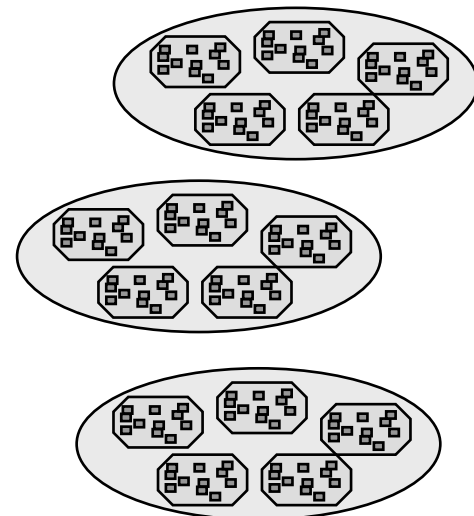
- 2D analog representation of the sign space
- Accurate representation of signs
- Possibility to represent both manual and non-manual components of signs ...
- but... it consists of 38,711 different glyphs



Organizzazione ufficiale dei glifi di SignWriting (SWoff)



Official organization of SignWriting glyphs (SWoff)



CATEGORIA
(da 1 a 7)

GRUPPO
(da 1 a 10)

SIMBOLO DI BASE
(da 1 a 44)

VARIAZIONE
(da 1 a 5)

RIEMPIMENTO
(da 1 a 6)

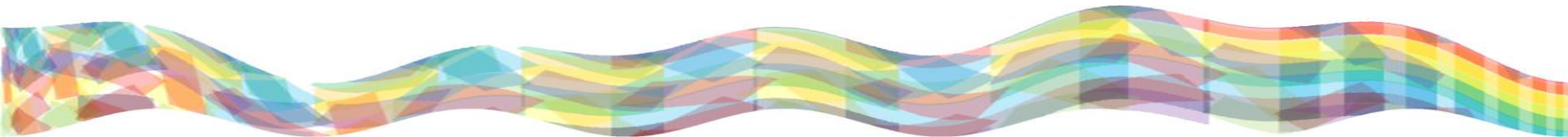
ROTAZIONE
(da 1 a 16)

Problemi di SignWriting

- Problemi:
 - Numero apparentemente esorbitante di glifi
 - Organizzazione dei glifi non trasparente (regole non esplicite)
- Conseguenze:
 - Scarsa diffusione del sistema
 - Paura della complessità da parte dei nuovi discenti
 - Problemi nella digitalizzazione del sistema
 - Problemi nell'uso di SW come strumento di ricerca in linguistica
- Soluzione:
 - Rendere l'organizzazione di SW più trasparente, in modo da ridurne la complessità
 - ma senza ridurre il numero di glifi (espressività)

Problems of SignWriting

- Problems:
 - Amount of glyphs apparently exorbitant
 - Non-transparent organization of glyphs (non explicit rules)
- Consequences:
 - Scarce diffusion of the system
 - Fear of complexity by new learners
 - Problems in system digitalization
 - Problems of using SW as a tool for linguistic research
- Solution:
 - Making the SW organization more transparent to reduce its complexity
 - yet without reducing the number of glyphs (expressiveness)



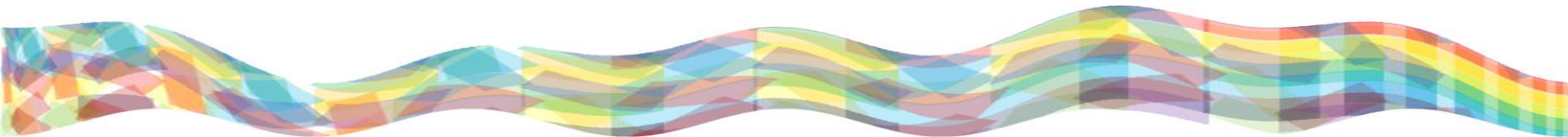
Ri-classificazione 1 (SW_{recl1})

- Per rendere più esplicite le regole, SW è stato riorganizzato basandosi sulle caratteristiche funzionali e grafiche dei singoli glifi
- individuazione di “glifi-prototipi” (GP)
- esplicitazione delle regole di passaggio tra GP e “glifi-varianti” (GV)

GP □ (1 di 242 1 of 242)	Mano destra / Right hand		
	Palmo / Palm	Taglio / Side	Dorso / Back
Piano verticale / Vertical plan			
Piano orizzontale / Horizontal plan			
GP □ (1 di 242 1 of 242)	Mano sinistra / Left hand		
	Palmo / Palm	Taglio / Side	Dorso / Back
Piano verticale / Vertical plan			
Piano orizzontale / Horizontal plan			

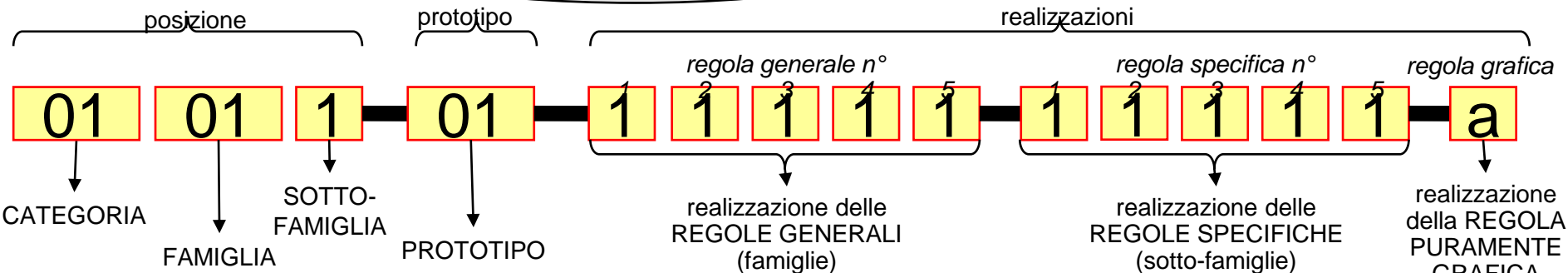
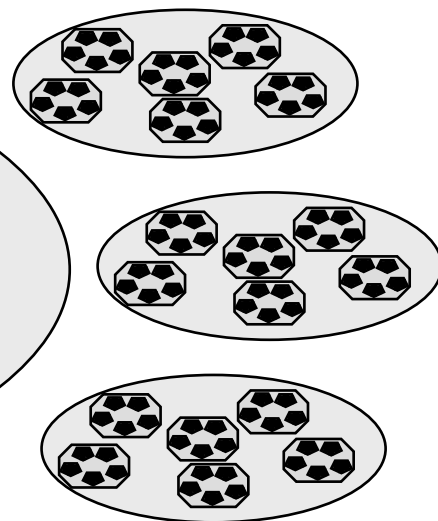
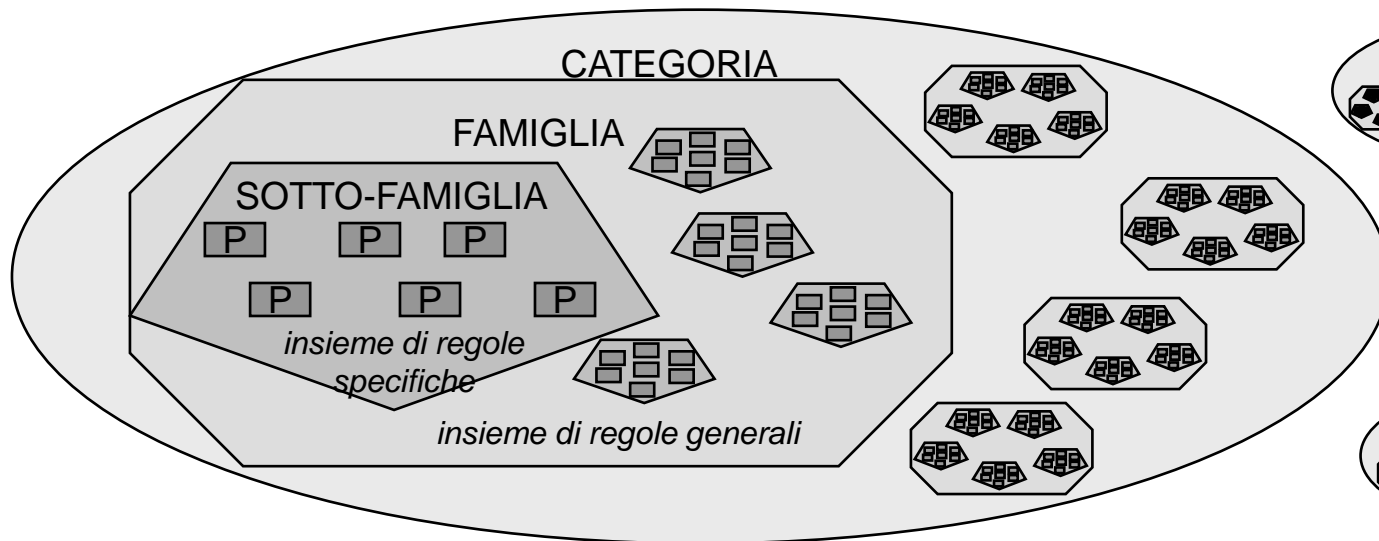
Re-classification 1 (SW_{recl1})

- In order to make rules more explicit, SW has been re-arranged according to the functional and graphical features of each glyph
- identification of “prototype-glyphs” (PGs)
- explication of transfer rules between PGs and “variant-glyphs” (VGs)



Ri-classificazione di SignWriting (SW_{recl1})

Re-classification of SignWriting (SW_{recl1})

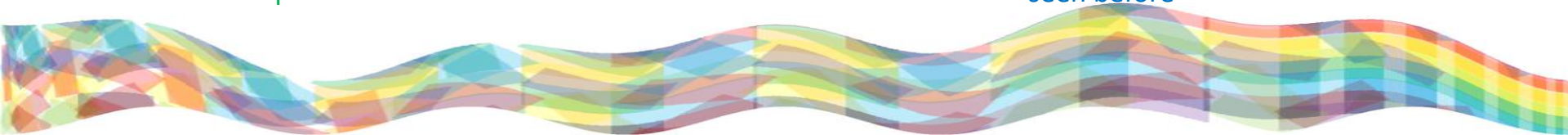


Applicazione didattica della SW_{recl1}

- Rendere più esplicite le regole permette un più facile apprendimento da parte di studenti sordi e udenti
- Test su studenti udenti:
 - 6 ore di apprendimento usando SW_{off}
 - produzione di vignette basiche
 - impossibilità di capire da soli il funzionamento di nuovi glifi
 - 6 ore di apprendimento usando SW_{recl1}
 - produzione di piccoli testi scritti e trascritti
 - possibilità di inferire le trasformazioni da GP a GV anche per glifi mai visti prima

Didactic application of SW_{recl1}

- Making rules more explicit allows for an easier learning by hearing or deaf students
- Test on hearing students:
 - 6 hours of learning using SW_{off}
 - production of basic cartoon
 - inability to understand without help how new glyphs operate
 - 6 hours of learning using SW_{recl1}
 - production of short written and transcribed texts
 - possibility to infer transformations from PGs to VGs even for glyphs not seen before

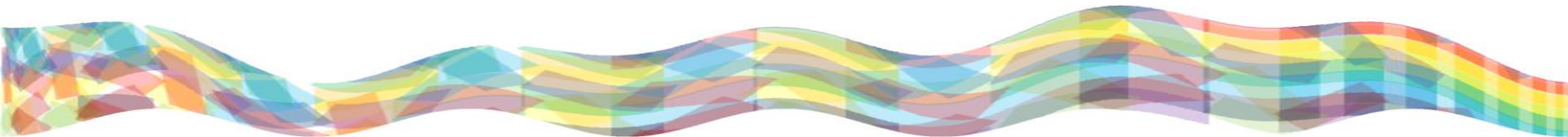


Applicazione informatica della SW_{recl1}: SWift

- SW_{recl1} mira anche a favorire l'informaticizzazione di SW per:
 - facilitare l'utilizzo di SW
 - migliorare l'utilizzo di SW come strumento di notazione linguistica
- Nuovo editor di testi in SW: SWift (SignWriting improved fast transcriber)
 - struttura e interfaccia "sordo-centrica"
 - strumento di ricerca dei glifi (pre-composizione) basato su SW_{recl1}

Computer application of SW_{recl1}: SWift

- SW_{recl1} aims also at supporting the computerization of SW to:
 - facilitate the use of SW
 - improve the use of SW as a linguistic notation tool
- A new text editor in SW: Swift (SignWriting improved fast transcriber)
 - "deaf-centered" interface and structure
 - glyph search tool (pre-composition) based on SW_{recl1}

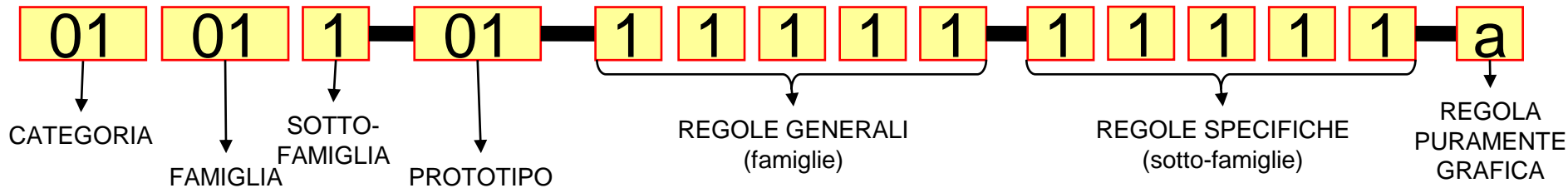


Applicazione scientifica della SW_{recl1}

- Grazie al nuovo codice di SW_{recl1}, è possibile identificare le singole caratteristiche di un glifo, facilitando l'analisi formale delle caratteristiche del segno
- ex: è possibile ricercare un particolare angolo di orientamento della mano, indipendentemente dalla sua configurazione o dal lato

Scientific application of SW_{recl1}

- Thanks to the new coding of SW_{recl1}, it is possible to identify the peculiar features of a glyph, therefore facilitating the formal analysis of a sign
- e.g., it is possible to search for a certain orientation angle of the hand, independently from its configuration or side

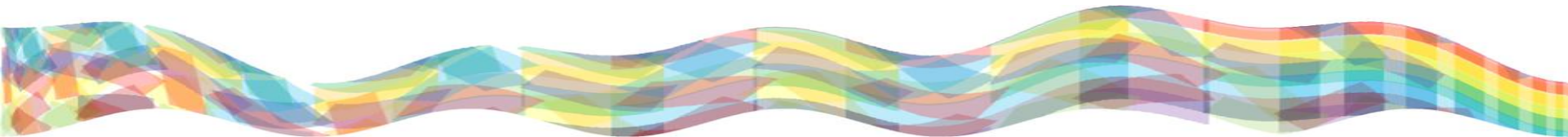


Dalla scrittura a mano alla scrittura digitale

- Gli utenti sordi campionati tendono a scrivere a mano e, solo dopo, a digitalizzare
- Idea: creare un OGR (Optical Glyph Recognition)
 - diverso dall'OCR "classico" per
 - numero di glifi di SW
 - tipo grafico di SW
 - disposizione 2D dei glifi di SW

From handwriting to digital writing

- The sampled deaf users tend to write by hand first, and only thereafter to digitalize
- Idea: creating an OGR (Optical Glyph Recognition)
 - different from the "classical" OCR by
 - amount of glyphs
 - graphic features of SW
 - 2D layout of glyphs

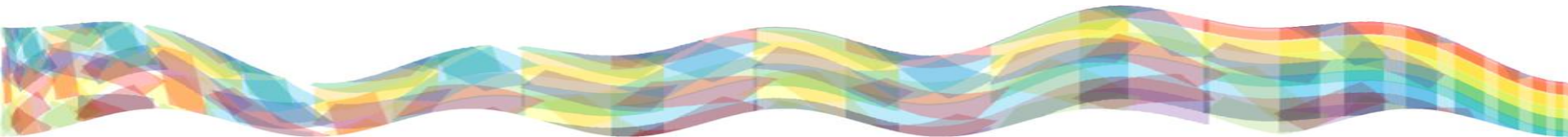


Ri-classificazione 2 (SW_{recl2})

- Necessità di riorganizzare SW nuovamente, per guidare meglio il riconoscimento:
 - parametri funzionali
 - parametri topologici
 - parametri geometrici
- Corrispondenza tra SW_{recl1} e SW_{recl2}

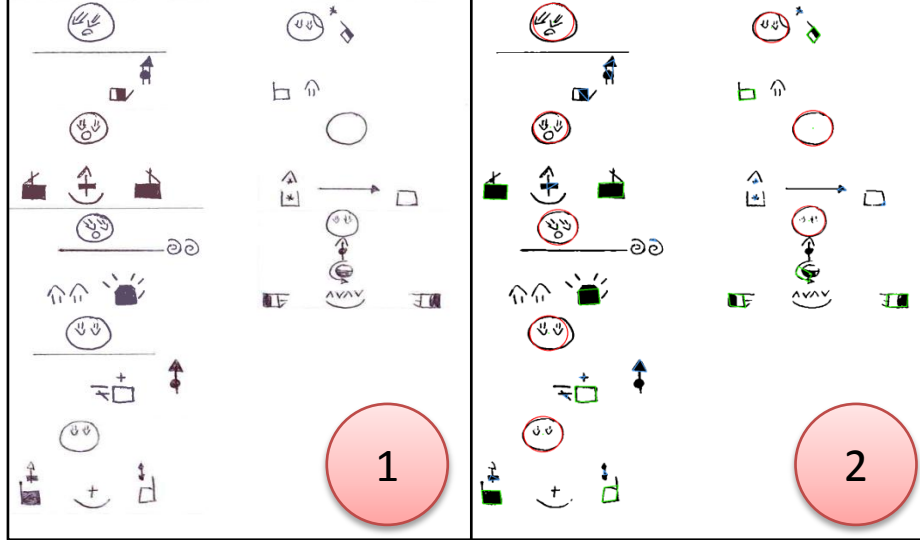
Re-classification 2 (SW_{recl2})

- Need to re-arrange SW once again, to better support recognition:
 - functional parameters
 - topological parameters
 - geometrical parameters
- Coherence between SW_{recl1} and SW_{recl2}



Scrittura Digitale: OGR

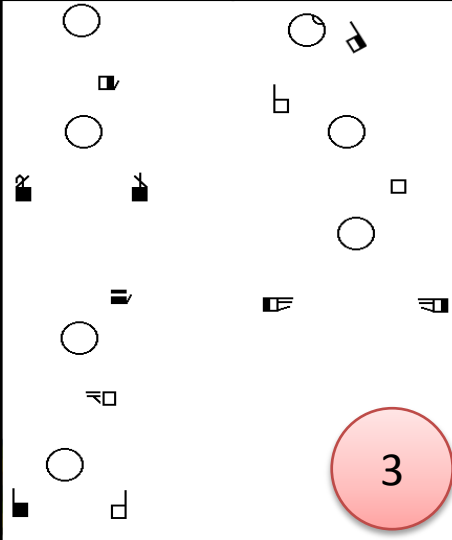
1. Elaborazione di un testo scritto a mano in SW
 - riconoscimento di testa (solo forma) e mani



Digital Writing: OGR

1. Processing of a handwritten SW text
 - recognizing head (shape only) and hands

2. Identificazione e elaborazione delle forme geometriche nel testo
3. Risultato dell'elaborazione OGR
 - i glifi sono stati digitalizzati



2. Identifying and processing any geometric shape within the text
3. OGR processing results
 - glyphs have been digitalized

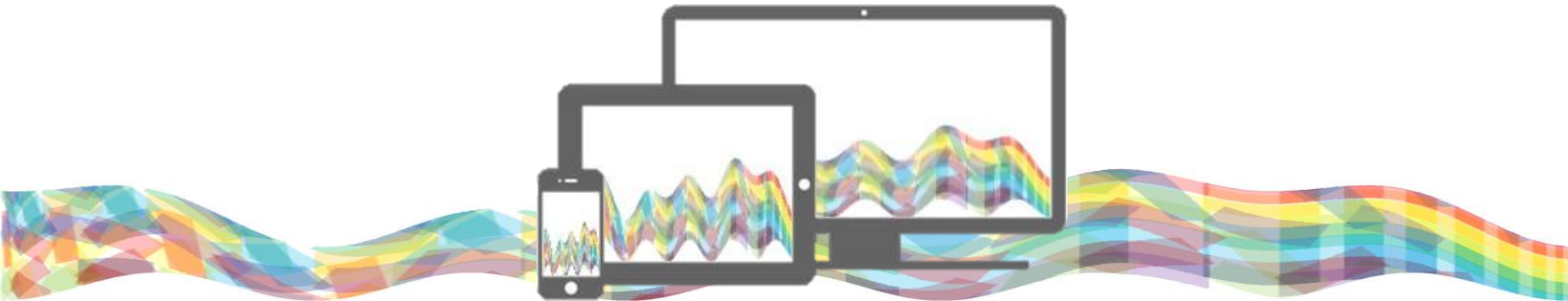


Scrittura Digitale - OGR

- OGR può essere applicato in (almeno) due scenari:
 - digitalizzazione di corpora cartacei in SW preesistenti
 - scrittura e digitalizzazione in tempo reale da dispositivi mobili o fissi

Digital Writing - OGR

- OGR can be employed in (at least) two scenarios:
 - digitalization of SW corpora already existent on paper
 - real-time writing and digitization via mobile or fixed devices

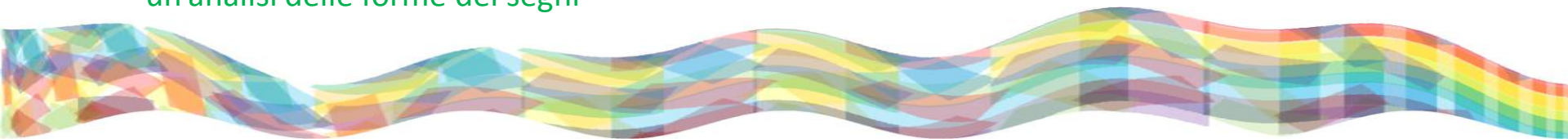


Conclusioni dirette

- SW_{rec1} e SW_{rec2} sono alla base di applicazioni per una facile digitalizzazione di documenti in SW. Questo consentirà:
 - la diffusione del sistema grafico presso la comunità sorda
 - nuove opportunità di analisi delle LS basandosi sulla “fedeltà” grafica di SW
 - resta da pensare l’integrazione con i maggiori software di annotazione di dati linguistici multimodali (ELAN, ANVIL, iLex ...)
- SW_{rec1} permette una maggiore comprensione dei sotto-parametri che entrano in gioco nella realizzazione di un segno
 - nuovo strumento di indagine linguistica per un’analisi delle forme dei segni

Direct conclusions

- SW_{rec1} and SW_{rec2} are the basis of applications for an easy digitalization of documents in SW. This will in turn allow:
 - the dissemination of the graphic system among the deaf community
 - new opportunities for the analysis of SLs based on the graphic “fidelity” of SW
 - it remains to consider an integration with leading software for multimodal annotation of linguistic data (ELAN, ANVIL, Ilex ...)
- SW_{rec1} allows a better understanding of the sub-parameters that come into play in the realization of a sign
 - new tool for linguistic analysis of the sign forms

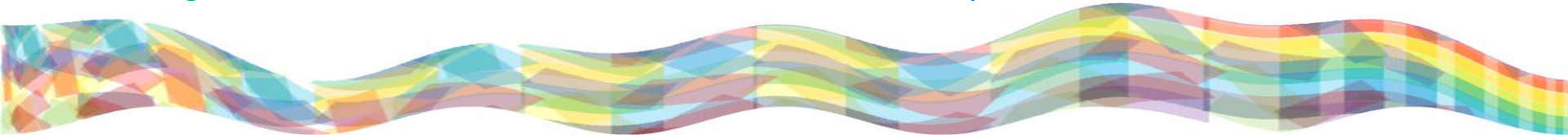


Conclusioni metodologiche

- Lo studio dettagliato (“in vitro” ma anche tramite utenti sordi) delle caratteristiche funzionali e grafiche di SW ha permesso di porre le basi per la creazione di un protocollo di analisi dei sistemi di rappresentazione grafica delle LS
 - miglioramento di SW?
 - studio di altri sistemi?
 - sviluppo di un sistema più performante di rappresentazione grafica delle LS?

Methodological conclusions

- The detailed study (“in vitro” but also by means of deaf users) regarding the functional and graphical features of SW allowed to build the basis to create a protocol to analyze the systems of graphical representation for SLs
 - SW enhancement?
 - study of other systems?
 - development of a more efficient system for the graphical representation of SLs?

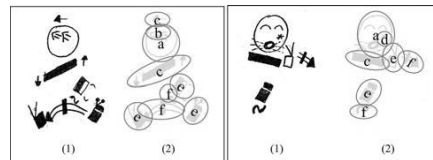


■ Digital Humanities 2014

Book of Abstracts

al contempo, non possono essere rappresentate adattando un sistema di scrittura nato per le lingue vocali (come l'alfabeto latino o anche l'alfabeto fonetico internazionale) a causa della differenza di canale comunicativo e della conseguente multi-linearità della lingua. Tale assenza provoca difficoltà nei compiti di notazione (ossia rappresentazione delle forme significanti), annotazione (ossia rappresentazione dei fenomeni linguistici) e successiva analisi di produzioni segnate.

Attualmente, la maggior parte delle analisi delle LS si basa sull'uso delle cosiddette "glosse", etichette verbali nella lingua vocale del ricercatore che permettono - secondo il loro scopo e il loro grado di complessità - di annotare il significato del segno, le caratteristiche della sua forma significante o anche altre informazioni che lo riguardano. Sebbene possa trattarsi di etichette anche molto precise, rimane difficile - se non impossibile - ricostruire la forma significante delle produzioni segnate attraverso queste etichette: si tratta quindi di soluzioni che permettono l'annotazione ma non la notazione delle LS. La soluzione usata, che aggira il problema della rappresentazione grafica della lingua, è quella di associare temporalmente a tali etichette il video originale, grazie all'uso di software di annotazione linguistica come ELAN¹ o ANVIL².



Segni rappresentati con SW (1) e specifica degli elementi descritti (2): (a) espressione facciale, (b) direzione dello sguardo, (c) movimento e posizione, della testa e del corpo, (d) contatto delle mani, (e) configurazione delle mani, (f) movimento e posizione delle mani e delle braccia

Tra i non molti sistemi esistenti per rappresentare graficamente le LS vi è SignWriting (SW) di V. Sutton³, nato per raffigurare le posizioni ed i movimenti del corpo (in particolare nella danza) e poi applicato alle LS. E' stato mostrato⁴ che esso rappresenta una valida soluzione in quanto consente sia la scrittura che la trascrizione di produzioni segnate, mantenendo le caratteristiche delle LS e permettendo una buona leggibilità dei testi così concepiti. Tale strumento può quindi essere utilizzato sia dai ricercatori a fini scientifici che dalla comunità sorda in generale.

Dal punto di vista del sistema grafico, SW è organizzato in 37 811 glifi (unità grafiche minime) che vengono disposti in una vignetta (spazio bidimensionale che rappresenta analogicamente lo spazio segnico tridimensionale) per comporre un segno [Fig. 1]. Ogni glifo rappresenta una componente manuale (configurazione, movimento, coordinazione, etc.) o non-manuale (espressione del viso, movimento del corpo, etc.) del segno. E' così possibile rappresentare sia le unità lessematiche⁵ che le strutture di grande iconicità⁶ sia segni isolati che segni inseriti in un discorso.

L'ostacolo maggiore incontrato da SW è la sua scarsa diffusione, dovuta in gran parte al numero esorbitante di glifi che lo compongono: tale quantità sarebbe, infatti, sufficiente a scoraggiare anche l'apprendente più volenteroso. E' stato tuttavia dimostrato^{4,7} che, una volta passata la diffidenza iniziale, è possibile - per soggetti segnanti - cominciare ad utilizzare il sistema in poche ore, grazie ad una forte regolarità di questi glifi. In ogni caso, il numero considerevole di glifi pone problemi anche alla digitalizzazione delle produzioni in SW ed ad un loro riconoscimento automatico, il che si riflette negativamente sulla possibilità concreta di usare SW a scopo di ricerca, per esempio per analizzare il legame tra la forma di un segno ed i fenomeni linguistici.

Il lavoro qui proposto presenta una radicale riorganizzazione di SW (d'ora in poi chiamata SW^{msi}) per raggruppare i glifi secondo le loro caratteristiche comuni e per mettere in evidenza quelle regole implicite che permettono un suo facile

apprendimento nonostante la quantità di glifi esistenti. A tale scopo si è provveduto a raccogliere i glifi sulla base della loro funzione (configurazione, movimento della mano, direzione dello sguardo...) per poi andare a individuare le caratteristiche che consentono di passare da un glifo prototipico (GP) ad un glifo variazione (GV). Ad esempio, per le configurazioni, sono stati individuati 242 GP e 4 regole di variazione con relativi set di possibilità (mano x 2 possibilità, piano x 2, lato x 3 e rotazione x 8) [Fig. 2]. I GP sono anche stati classificati in modo da evidenziarne le somiglianze, e così, imparando le 4 regole e circa una trentina dei 242 GP, è possibile sfruttare 23 232 GV ossia oltre la metà dei glifi esistenti. La facilità di apprendimento di SW dopo la riorganizzazione con SW^{msi} è stata preliminarmente testata su un gruppo di studenti udenti di LS: i risultati sono incoraggianti in quanto, dopo appena 6 ore di formazione, i soggetti erano in grado di orientarsi facilmente all'interno di SW^{msi} (cosa non possibile con la vecchia classificazione) e di comporre e leggere brevi testi scritti in SW.

G _p	Mano destra			Mano sinistra		
	Palmo	Taglio	Dorso	Palmo	Taglio	Dorso
Piano verticale						
Piano orizzontale						

Ogni glifo prototipo che identifica una configurazione può, per trasformazione, dar vita a 96 glifi variazione basati sulla mano utilizzata e il suo orientamento

Oltre alle implicazioni nell'apprendimento di SW, SW^{msi} mira anche a favorire la sua informatizzazione, sia per permetterne un più facile utilizzo da parte degli utenti sordi e udenti, che per renderlo uno strumento utilizzabile nella ricerca linguistica. A tale scopo è nata la collaborazione tra gli autori - un linguista e due informatici - con l'idea di cercare di risolvere, un passo alla volta, le difficoltà legate alla digitalizzazione del sistema grafico.

In primo luogo, è stato creato SWift (SignWriting Improved Fast Transcriber), un editor di testi in SW, che permette attraverso una struttura ed un'interfaccia "deaf-centered" (ossia interamente pensata e realizzata in collaborazione con gli utenti sordi) di produrre testi in SW. In SWift l'utilizzo di SW^{msi} ha permesso di rendere più rapido, efficace e "user friendly" lo strumento di ricerca dei glifi che verranno usati dall'utente per comporre il proprio testo.

Nel corso delle loro ricerche su SW, gli autori hanno notato la tendenza degli utenti osservati a comporre prima a mano e poi a ricopiare al computer i testi: ciò ha portato alla necessità di implementare applicazioni di OGR (Optical Glyph Recognition), con caratteristiche diverse dal classico OCR (Optical Character Recognition) in quanto più vicine al riconoscimento di scritture di tipo ideografico. In effetti, OGR deve prendere in considerazione un numero molto elevato di glifi, organizzati in uno spazio grafico non lineare e disposti secondo regole di composizione non rigide. Una volta concluso il suo sviluppo, OGR effettuerà il riconoscimento basandosi esclusivamente sulle caratteristiche geometriche e topologiche dei glifi: si è pertanto resa necessaria - in fase di progettazione - una seconda fase di riclassificazione di SW (SW^{ms2}) basata su queste caratteristiche.

Le applicazioni menzionate, basate su SW^{msi} e SW^{ms2}, dovrebbero permettere una facile digitalizzazione di documenti prodotti in SW, il che favorirà la diffusione del sistema grafico presso la comunità sorda. Tuttavia, ciò che preme ai linguisti che usano SW è che esso possa finalmente diventare uno strumento "utilizzabile" per fare analisi sulle forme della LS. La poca integrazione, fino ad oggi, tra SW ed i software di annotazione linguistica più comunemente usati implica, infatti, una seria difficoltà per sfruttare le potenzialità di SW

come sistema di notazione delle LS. Integrando SW con ELAN, sarebbe possibile mettere in evidenza le relazioni tra significante e significato, le correlazioni tra forme del segno e strutture linguistiche, ecc. ossia tra la notazione e l'annotazione delle LS.

Oltre alla sfida dell'integrazione informatica, si pone un duplice problema legato alla natura stessa di SW: visto il numero elevato di glifi e la scarsa standardizzazione di utilizzo del sistema grafico, è difficile realizzare trascrizioni che permettano poi un effettivo confronto sulle forme. Tale problema è però superabile grazie alla SW^{msi} che - mettendo in evidenza le caratteristiche intrinseche ai singoli glifi (caratteristiche trasversali a diversi GP) - permette di collegare tra di loro realizzazioni grafiche diverse di forme segniche simili, o anche caratteristiche identiche in forme segniche diverse.

SW^{msi} ed SW^{ms2} si pongono quindi così come primi passi per un'informatizzazione di SW che dovrebbe permettere lo sviluppo e la diffusione del sistema sia all'interno della comunità sorda che nell'ambito della ricerca linguistica sulla LS.

References

1. **ELAN Team.** 2002. *ELAN - Electronic Linguistic Annotator*. Nijmegen (Netherlands), Max Planck Institute of Psycholinguistics. www.lat-mpi.eu/tools/elan/ [software, disponibile on line].
2. **Kipp M.** 2000. *ANVIL. University of Applied Sciences, Augsburg*. www.anvil.software.de/index.html [software, disponibile on line].
3. **Sutton V.** 1995. *Lessons in SignWriting: textbook & workbook*. Deaf Action Committee for Sign Writing, La Jolla (CA).
4. **Bianchini C.S.** 2012. *Analyse métalinguistique de l'émergence d'un système d'écriture des Langues des Signes: SignWriting et son application à la Langue des Signes Italienne (LIS)*. Tesi di dottorato, Université Paris VIII & Università degli Studi di Perugia, 512 p.
5. **Cuxac C. & Antinoro Pizzuto E.** (2010). *Émergence, norme et variation dans les langues des signes: vers une redéfinition notionnelle*. in: Garcia B., Derycke M. (eds) "Sourds et langues des signes: norme et variations". Langage et société, 131: 37-53.
6. **ELAN Team** (2002). *ELAN - Electronic Linguistic Annotator*. Nijmegen (Netherlands), Max Planck Institute of Psycholinguistics. www.lat-mpi.eu/tools/elan/ [software, disponibile on line].
7. **Borgia F., Bianchini C.S., Dalle P. & De Marsico M.** (2012). *Resource production of written forms of Sign Languages by a user centered editor, SWift (SignWriting improved fast transcriber)*. Proc. VIII Int. Conf. LREC: 3779-3784.

Riorganizzare SignWriting per favorire la ricerca linguistica sulle Lingue dei Segni

Bianchini, Claudia S.
claudia.savina.bianchini@univ-poitiers.fr
 EA3816 FoReLL - Université de Poitiers

Borgia, Fabrizio
fabrizio.borgia@uniroma1.it
 "Sapienza" Università di Roma

De Marsico, Maria
demarsico@di.uniroma1.it
 "Sapienza" Università di Roma

Le Lingue dei Segni (LS) sono lingue utilizzate dalla maggior parte dei sordi del mondo per comunicare tra di loro; si tratta di lingue storico-naturali caratterizzate dall'uso del canale visuo-gestuale. La loro rappresentazione grafica è una sfida per la ricerca linguistica in quanto le LS non hanno ancora sviluppato un sistema di scrittura condiviso dalla comunità e,