



HAL
open science

Représenter les langues des signes sous forme écrite : questions qui ont besoin (encore aujourd'hui) d'être posées

Claudia S. Bianchini

► **To cite this version:**

Claudia S. Bianchini. Représenter les langues des signes sous forme écrite : questions qui ont besoin (encore aujourd'hui) d'être posées. CORELA - COgnition, REprésentation, LAngage, 2021, 19 (2), pp.1-37. 10.4000/corela.13190 . hal-03740978

HAL Id: hal-03740978

<https://hal.science/hal-03740978v1>

Submitted on 30 Jul 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0
International License

Représenter les langues des signes sous forme écrite : questions qui ont besoin (encore aujourd'hui) d'être posées

Claudia S. Bianchini

Hommage à Elena Antinoro Pizzuto à 10 ans de son décès

« The next day, Monday, we were playing in the fields and this boy said to me, “See that bird standing on the stump there? What’s the name of it?”. I said, “I haven’t got the slightest idea.” He said, “It’s a brown-throated thrush. Your father doesn’t teach you much about science.” I smiled to myself, because my father had already taught me that [the name] doesn’t tell me anything about the bird. He taught me “See that bird? It’s a brown-throated thrush, but in Germany it’s called a halsenflugel, and in Chinese they call it a chung ling and even if you know all those names for it, you still know nothing about the bird – you only know something about people; what they call that bird. Now that thrush sings, and teaches its young to fly, and flies so many miles away during the summer across the country, and nobody knows how it finds its way,” and so forth. There is a difference between the name of the thing and what goes on. »

(R. Feynman, 1969. In *The Physics Teacher*, 7(6); p. 313-320)

Introduction¹

- 1 La quasi-totalité des langues ayant existé depuis la nuit des temps n'ont pu laisser de traces car elles étaient dépourvues d'un système d'écriture ; aujourd'hui encore, la majorité des langues du monde sont purement orales, non-écrites (Coulmas, 2013). Toutefois, si l'exigence se présente, il est possible de représenter graphiquement (presque) toutes les langues orales en adaptant un système phonographique préexistant. L'alphabet phonétique international (API) est la manifestation la plus évidente de cette possibilité, mais il serait aussi possible de réadapter un système alphabétique ou syllabique utilisé pour une autre langue écrite ou même d'inventer une nouvelle série de caractères et de règles de composition reposant sur les principes de fonctionnement d'une écriture phonographique existante. Cela est possible pour toutes les langues constituées de séquences de sons, c'est-à-dire pour toutes les langues vocales (LV).
- 2 Bien qu'étant un système graphique, l'API n'est toutefois pas considéré ici comme un système écriture. En effet, la notion d'écriture peut avoir des acceptions diverses selon les points de vue adoptés (Meletis, 2020), mais l'idée retenue ici est qu'il s'agit d'un système de représentation d'une langue spécifique à travers un ensemble de symboles (caractères) conventionnels, inscrits sur un support, en appliquant des règles conventionnelles². Ces caractères et ces règles graphiques constituent des savoirs partagés par les scripteurs, qui doivent aussi être membres de la communauté des locuteurs de la langue en question³, car l'écriture sans rapport à la langue est juste un ensemble de symboles à déchiffrer. Tout comme la langue en elle-même, l'écriture reflète — tout en l'influençant — la langue, la culture, les croyances, l'imaginaire et le rapport au monde des personnes qui la pratiquent, même lorsque ce lien semble avoir été estompé par la banalité de la pratique de l'écrit. C'est ce lien culturel qui fait que l'API, bien qu'en mesure de représenter graphiquement toute les LV du monde, ne peut — aux yeux de l'auteur de cet article — être considéré comme un système d'écriture, mais seulement comme un système de représentation graphique utilisé pour transcrire la forme signifiante d'une langue, autrement dit un système de transcription.
- 3 Si toutes les LV peuvent être transcrites avec l'API, il n'en est pas de même pour les langues des signes⁴ (LS), utilisées par de nombreux sourds (surtout sourds profonds, mais non uniquement) pour communiquer entre eux et avec leur entourage⁵. S'agissant de langues visuo-gestuelles, elles ne partagent pas le canal audio-phonatoire avec les LV, ce qui rend impossible l'utilisation d'un système phonographique pour les représenter. Bien que ces langues existent probablement depuis la naissance du premier être humain sourd, aucune d'entre elle ne semble avoir laissé une trace écrite de son passage : il est donc possible d'affirmer qu'actuellement les LS sont toutes des langues orales et non-scriptibles, du moins à travers l'adaptation d'une typologie d'écriture existante⁶. Si quelques tentatives d'écrire les LS ont vu le jour (notamment SignWriting [Sutton, 1995], cf. §4), aucune de ces propositions graphiques n'a réussi à construire le lien entre la langue, l'écriture et la culture des locuteurs et elles sont restées cantonnées à de petites communautés d'initiés (Hopkins, 2008). Il s'ensuit que, bien que leurs inventeurs déclarent vouloir donner la possibilité d'écrire les LS, aucun

de ces systèmes ne peut être considéré, à l'heure actuelle, comme un véritable système d'écriture des LS.

- 4 L'absence d'une solution pour écrire les LS a des retombées négatives non seulement sur la possibilité de transmettre la langue et la culture Sourde⁷, mais aussi sur la reconnaissance des LS et des droits linguistiques de ses locuteurs. Depuis la constitution des premières écoles pour sourds en France, au XVIII^e siècle, les LS ont connu une histoire fluctuant entre la reconnaissance et le déni, couronnée— après le Congrès de Milan en 1880 — par l'interdiction d'utiliser les signes dans l'éducation des sourds (Encrevé, 2012). Au bout de près de 100 ans d'oubli pédagogique, social et aussi linguistique, des mouvements sociaux de « Réveil Sourd » ont conduit à la revitalisation de plusieurs LS⁸. Il demeure qu'aujourd'hui, beaucoup de LS sont encore considérées comme des langues menacées de disparition (Bickford et McKay-Cody, 2018). Or, dans sa Foire aux Questions sur les langues en péril, à la question « *que peut-on faire pour empêcher une langue de disparaître ?* » l'UNESCO (s.d.) répond « *La chose la plus importante [...] est de créer des conditions favorables pour que ses locuteurs la parlent et l'enseignent à leurs enfants. Cela nécessite [entre autres] une collaboration créative entre les membres de la communauté et les linguistes afin d'élaborer un système d'écriture [...]* ».
- 5 L'absence d'écriture n'est pas seulement un obstacle à la reconnaissance sociale des LS, elle est également lourde de conséquences pour les études que les linguistes mènent sur ces langues. C'est l'avis de Elena Antinoro Pizzuto (1952-2011†) (EAP⁹), chercheuse de l'Istituto di Scienze e Tecnologia della Cognizione del Consiglio Nazionale delle Ricerche (ISTC-CNR) de Rome qui, entre 2000 et 2011, a centré ses recherches sur la représentation graphique des LS. Adhérant au Modèle Sémiologique proposé par Christian Cuxac (2000 ; cf. §1), EAP étudie la LS Italienne (LIS) en se fondant sur l'analyse de corpus discursifs et en centrant son intérêt sur le rôle de l'iconicité dans la production du sens. À ces concepts, EAP rajoute toutefois l'idée qu'il est impossible de faire de la recherche linguistique sur les LS sans s'interroger sur leur forme et donc sur leur représentation graphique. Elle publie ainsi, entre 2000 et 2010, des articles aux titres éloquentes : *The notation of signed texts: open questions and indications for further research* (Pizzuto et Pietrandrea, 2001) ; *Representing Signed Languages in written form: questions that need to be posed*¹⁰ (Pizzuto et al., 2006) ; *The representation issue and its multifaceted aspects in constructing Sign Language corpora: questions, answers, further problems* (Antinoro Pizzuto et al., 2008) ; *Representing Signed Languages: theoretical, methodological and practical uses* (Antinoro Pizzuto et al., 2010).
- 6 EAP affirme que la description de la signification d'un signe, de la direction du regard ou du mouvement de la main, ne peut pas être reléguée à une simple étiquette d'annotation dans la LV du chercheur (Antinoro Pizzuto, 2009). Il faut, selon elle, que ces éléments soient représentés par un système graphique qui représente la LS en question et qui permette de voir de façon synthétique toutes les informations sur la forme des signes, afin de réussir à décrypter la relation entre la forme et le sens en LS. Son but est de fournir un système qui permette enfin d'abandonner une recherche linguistique où l'analyse des données est remplacée par l'analyse d'étiquettes descriptives, fruit de l'interprétation des chercheurs.
- 7 Toutefois, pour EAP, la mise au point d'un système visant à représenter graphiquement les LS n'a pas seulement pour but d'améliorer la connaissance des LS, mais doit servir à modifier les représentations de la culture scientifique qui voient les sourds uniquement comme des individus participant passivement à la recherche. Doter les LS d'une

écriture doit permettre aux sourds d'avoir accès aux données sur leur langue à travers leur propre langue, leur permettant ainsi d'assumer un rôle de chercheur dans les analyses sur ces corpus. Les difficultés rencontrées par les sourds dans la maîtrise de la LV écrite de leur pays (Perini, 2013) peuvent limiter leur accès aux phases actives de la recherche en Science du Langage, mais aussi à la recherche scientifique en général. Résumant sa pensée par la phrase « faire de la recherche avec les sourds et non sur les sourds », EAP rassemble autour d'elle, entre 2000 et 2011, une équipe composée d'ingénieurs d'étude sourds (qu'elle forme elle-même à la linguistique et au Modèle Sémiologique) et de quelques « apprentis-chercheurs » (en thèse ou en master), sourds et entendants.

- 8 En janvier 2011, le décès de EAP laisse un grand vide du point de vue humain, mais aussi un grand nombre de questions de recherche encore en suspens. Cet article, écrit 10 ans plus tard par une des apprenties-chercheuses qu'elle avait formées, s'interroge sur l'actualité des questions posées par EAP à la lumière des progrès technologiques qui ont marqué la recherche sur les LS ces dernières années, et sur les solutions qu'il est possible d'envisager aujourd'hui, dans une optique pluridisciplinaire.

1. La question graphique dans l'histoire de la linguistique des LS

- 9 L'histoire de la représentation graphique des LS, vue comme moyen de figer les caractéristiques linguistiques de ces langues, est relativement récente. Elle est étroitement liée à l'histoire des LS et de l'éducation des sourds d'un côté et à l'histoire des courants linguistiques étudiant ces langues de l'autre.
- 10 À partir de la moitié du XVIII^e siècle et pendant environ 100 ans, on assiste à l'Âge d'or des LS, caractérisé par l'essor de méthodes d'éducation des sourds basées sur les LS (notamment grâce à la méthode signée de l'Abbé de l'Épée). Cette période est caractérisée par la création de dictionnaires de LS (Bonnal-Vergès, 2008), qui se fondent sur le dessin des signes ou sur la description, en LV, de la forme du signe lui-même ou de l'idée que cette forme est supposée représenter iconiquement. La première tentative de mise au point d'un système graphique propre aux LS semble être la Mimographie de Bébien (1825). Après avoir subdivisé la LS Française (LSF) en paramètres (forme, orientation, emplacement et dynamiques des mains d'un côté, mimique faciale de l'autre), Bébien associe un symbole graphique à chacune des formes qu'il identifie ; disposés linéairement dans un ordre prédéfini, ces symboles permettent la description de signes isolés de leur contexte. Le Congrès de Milan de 1880, en interdisant l'utilisation scolaire des LS, mit toutefois fin à toute tentative de recueillir et d'analyser les LS, plongeant ainsi les LS dans l'oubli.
- 11 En 1960, un chercheur américain, William Stokoe, décide de réhabiliter la LS américaine (ASL) comme vraie langue à travers, entre autres, la constitution d'un dictionnaire d'ASL et la démonstration que les LS sont doublement articulées (Stokoe, 1960). Pour réaliser ces deux branches de sa recherche, il crée un système basique de représentation graphique des différents « chérèmes » (c.-à-d. l'équivalent gestuel des phonèmes), qu'il met en évidence dans des signes lexicaux exprimés en forme citationnelle. Son système, bien que graphiquement distinct et réalisé de façon scientifiquement plus rigoureuse, reprend les principes de la Mimographie : subdivision de l'ASL en paramètres (configuration, orientation, emplacement et

mouvement des mains, mais sans aucune référence aux expressions faciales), identifications d'environ 60 chérèmes distincts et association de chacun d'eux à un symbole graphique (une lettre, un chiffre ou un symbole présent dans un clavier de machine à écrire), et enfin disposition monolinéaire des symboles dans un ordre prédéterminé. Ce modèle sera ensuite repris par la majorité des chercheurs — assez rares (cf. § 4) — proposant un système de transcription des LS¹¹.

- 12 À partir des années 70, aux USA, avec l'essor du modèle générativiste de Noam Chomsky (1957), se développe une approche « assimilatrice » (Cuxac et Antinoro Pizzuto, 2010) à l'étude des LS. Le but des chercheurs générativistes est de confirmer l'existence d'une grammaire universelle à travers l'étude d'une langue dont les manifestations superficielles, caractérisées par l'utilisation du canal visuo-gestuel, sont radicalement différentes des LV. Ce sont donc les points communs entre LS et LV qui sont mis en exergue. Malgré la relation iconique évidente entre les signes et leurs référents, les premières recherches sur les LS s'attardent à démontrer que les signes (souvent étudiés dans leur forme citationnelle) tendent à perdre leur iconicité (Frishberg, 1975) et que les « vrais signes » sont arbitraires (Klima et Bellugi, 1988). Pour cette ligne de recherche, la substance corporelle des LS se révèle plus un obstacle qu'un atout, elle est donc très peu prise en compte, voire totalement ignorée, d'où un manque d'intérêt généralisé pour le problème de la représentation graphique des LS (Chevrefils *et al.*, 2021). Cette approche mène à la généralisation de l'usage des « gloses », c'est-à-dire que les signes sont représentés par des étiquettes d'annotation en LV, qui véhiculent le sens des signes sans en faire transparaître la forme.
- 13 En France, à partir de 1975, un mouvement social pro-LS, connu sous le nom de « Réveil Sourd » (Minguy, 2009), s'accompagne des premières recherches sur la LSF, menées notamment par Christian Cuxac et son équipe (voire, parmi d'autres : Cuxac, 1983, 2000 ; Sallandre, 2003) à l'Université de Paris 8. N'adhérant pas aux théories de Chomsky, ces chercheurs proposent un modèle « non assimilateur », c'est-à-dire qu'ils défendent une approche basée sur la recherche des caractéristiques spécifiques aux LS. L'iconicité, destinée à disparaître selon les générativistes, devient ici le principal facteur explicatif du fonctionnement des LS. Parmi les nombreuses intuitions de ce groupe de recherche (Cuxac et Antinoro Pizzuto 2010 ; Cuxac, 2013 ; Sallandre 2014), la plus importante aux fins du présent article est celle qui consiste à essayer de lier les différentes fonctions de la LSF à ses manifestations corporelles, à travers l'analyse de corpus multimodaux de narrations en LS : ce ne sont plus des signes en forme citationnelle ou des phrases élicitées à partir de phrases en LV qui sont étudiées, mais des discours spontanés en LS. Cette démarche permet de mettre en avant non seulement le rôle des paramètres manuels identifiés par Stokoe (forme, emplacement, orientation et mouvement des mains) mais aussi certains paramètres non manuels, comme le regard, l'expression faciale et les mouvements du buste. Toutefois, ce groupe de recherche ne développe pas un système de représentation graphique des LS¹² : le besoin de représenter la forme signifiante de la LSF est satisfait par l'utilisation d'annotations en partitions, où chaque paramètre est codé sur une ligne à travers le recours à des étiquettes décrivant, en LV, la forme du paramètre jugé pertinente pour l'analyse.
- 14 Bien qu'adhérant à l'approche théorique de Cuxac, EAP refuse le recours aux étiquettes verbales¹³, que celles-ci ne représentent que le signifié ou qu'elles soient agrémentées d'informations sur la forme signifiante. Elle considère en effet que ces étiquettes

constituent un biais dans l'analyse, qui éloigne les chercheurs de la source de leurs données et qui risque de fournir une image faussée de la langue analysée.

2. Les « gloses » comme sources de biais dans l'analyse linguistique

- 15 Afin d'expliquer cette prise de position « anti-étiquettes verbales », EAP compare les étapes de création d'un corpus oral de LV et d'un corpus de LS. Quelle que soit la langue étudiée, les chercheurs, une fois le corpus constitué, doivent donner une forme aux données avant d'enrichir le corpus avec des annotations.
- 16 EAP se focalise sur la mise en forme graphique du matériel recueilli – c'est-à-dire (pour un corpus de LV orale) le passage d'un flux audio à un texte écrit – et sur le lien que celle-ci entretient avec les analyses. Afin d'exploiter les données du corpus, il faut établir des critères de segmentation et définir un système de transcription (qui peut être orthographique ou phonétique et aussi plus ou moins strict, selon les besoins de la recherche en cours). Ce système représentera la forme – la substance – de ce qui a été enregistré, sachant que les choix faits à ce stade influencent grandement les analyses menées sur le corpus.
- 17 Une fois le corpus mis en forme, il est possible de l'annoter, c'est-à-dire de l'enrichir d'étiquettes permettant d'identifier, analyser et comptabiliser les phénomènes langagiers retenus par les chercheurs. Malgré tous les efforts qu'il est possible de faire pour rester objectif, cet étiquetage est toujours influencé par les exigences et le point de vue scientifique du chercheur. Comme le résumant bien Hovy et Lavid (2010 ; p. 13) « corpus annotation [is] the process of enriching a corpus by adding linguistic and other information [...] in service of a theoretical or practical goal ».
- 18 Pour illustrer en quoi les étiquettes verbales constituent un biais pour l'analyse des LS, il est possible de prendre comme exemple un chercheur anglophone, qui analyse une vidéo dans laquelle un parlant dit « le chat poursuivait le chien » en *Romanesco*, la variante romaine de l'Italien (Tab.1).

(1a) SOURCE AUDIO	<i>non disponible sur un document imprimé</i>					
(1b) TRANSCRIPTION phonétique	er	gatto	insegwiva	er	kane	
(1b') TRANSCRIPTION standardisée	#il	gatto	inseguiva	#il	cane	
(1c) TRADUCTION globale	the cat was chasing the dog					
(1c') TRADUCTION mot à mot	the	cat	was chasing	the	dog	
(1d) GLOSES	DET.	CAT	TO CHASE	DET.	DOG	
(1e) ANNOTATIONS	D ms	N ms	V 3° s	D ms	N ms	

Tab. - Niveaux d'analyse pour la phrase en Romanesco (var. romaine de l'Italien) « le chat poursuivait le chien ».

En gris, les lignes qui seraient absentes

D/DET.=déterminant ; N=nom ; V=verbe ; m=masculin ; s=singulier ; 3°=troisième personne

indique un mot dont la forme diffère fortement entre l'italien standard et le Romanesco

Réinterprétation d'exemples fournis par Pizzuto et Pietrandrea (2001)

- 19 Faute de pouvoir constamment se référer aux données orales brutes – ce qui impliquerait de réécouter constamment l'intégralité du corpus –, le chercheur réalise une transcription phonétique (par exemple, mais pas nécessairement, en API) du signal sonore (1b) et/ou une transcription orthographique basée sur l'italien standard (1b'). Cette transcription lui permet (à lui ou même à un autre chercheur accédant à ces données) d'avoir une idée, avec une bonne approximation, de la forme originale et du

contenu de la phrase prononcée dans la bande sonore et de mener ses analyses sur cette approximation. Il peut aussi traduire du sens général de la phrase (1c) et celle de chaque unité (1c'), la traduction étant une opération de passage d'une langue à une autre introduisant une certaine marge de discordance. La traduction correspondra (plus ou moins, selon les cas) aux étiquettes (1d) choisies par le chercheur pour identifier les unités de sens qu'il a repérées, c'est-à-dire à ce qu'on appelle les « gloses ». Ayant à sa disposition toutes les lignes, le chercheur peut constamment vérifier la pertinence et la cohérence de ses choix de segmentation, de transcription et d'identification des unités de sens ; ces mêmes lignes permettront la vérification des données par les pairs, même dans le cas où ceux-ci ne peuvent pas accéder aux données brutes. La dernière ligne (normalement plus complexe que celle fournie ici) contient les annotations réalisées par le chercheur, qui pourront là encore être vérifiées par le lecteur en les mettant en relation avec la transcription de la source.

- 20 Si, au lieu d'être formulée en Romanesco, la phrase avait été signée dans la variante romaine de la LIS, le corpus subirait un traitement graphique très différent (Tab.2).

(2a) SOURCE VIDÉO	<i>non disponible sur un document imprimé</i>		
(2b) TRANSCRIPTION phonétique	<i>non disponible par manque d'un système de notation</i>		
(2b') TRANSCRIPTION standardisée	<i>non disponible par manque d'un système de notation</i>		
(2c) TRADUCTION globale	the cat was chasing the dog		
(2c') TRADUCTION mot à mot	DOG	CAT	CHASE
(2d) GLOSES/ÉTIQUETTES simples	DOG	CAT	CHASE
(2d') GLOSES/ÉTIQUETTES linéaires	[_{LH} DOG ^X][_{LH} PROF ^{sA} & _{RH} CAT ^X][_{LH} PROF ^{sA} & _{RH} CLS ^B][_{LH} RUN ^{A>C} & _{RH} CHASE ^{B>C}]		
(2d'') GLOSES/ÉTIQUETTES multilinéaire	LII RH	[DOG ^X]	[PROF ^{sA}] [PROF ^s RUN ^{A>C}] [CAT ^X] [PROF ^{sB}] [PROF ^s CHASE ^{B>C}]
(2e) ANNOTATIONS	N s	N s	V 3° s

Tab. - Niveaux d'analyse pour la phrase en langue des signes italienne (var. Rome) "le chat poursuivait le chien"

En gris, les lignes qui seraient absentes

Légende : N=nom ; V=verbe ; s=singulier ; 3°=troisième personne

_{LH}=main gauche ; _{RH}=main droite ; ^{X,A,B,C}=emplacements ; PROF^s=proforme de forme S ; _{LH}&_{RH}=coarticulation

Réinterprétation d'exemples fournis par Pizzuto et Pietrandrea (2001)

- 21 Dans ce script, la transcription (2b et 2b'), qui devrait servir de lien entre les données brutes et les analyses, est absente. En effet, non seulement les LS sont dépourvues d'un système d'écriture mais, en tant que langues non vocales, elles ne peuvent pas être représentées par les systèmes de transcription phonétique habituels. De ce fait, ce script qui devrait représenter une phrase en LIS ne renferme pas un seul élément en LIS. La LIS y est remplacée par les « gloses », qui ne devraient être rien de plus que des commentaires complétant les transcriptions, alors qu'ici elles en deviennent les substituts. L'utilisation des « gloses » comme base de l'analyse éloigne celle-ci des données brutes et peut engendrer de graves biais, dus au passage d'une langue à l'autre : par exemple, gloser un signe par un adjectif en LV *ne signifie pas obligatoirement que ce signe soit un adjectif en LS* ; gloser deux signes par le même mot en LV *ne signifie pas obligatoirement qu'il s'agisse du même signe à l'origine* ; gloser deux signes synonymes par deux mots synonymes en LV *ne signifie pas obligatoirement que la nuance*

entre les deux synonymes en LS correspond à celle en LV ; etc. Une étiquette « *ne signifie pas obligatoirement que* » le phénomène linguistique correspond à celui qui est présent sur l'étiquette, mais la nature de l'étiquette en LV risque tout de même d'influencer les jugements et les analyses des chercheurs sur les LS, surtout si ceux-ci n'ont jamais eu accès aux données brutes. Ces problèmes peuvent émerger à chaque fois qu'une langue est analysée à travers une autre langue, mais dans le cas des LV la ligne de transcription constitue un lien indissoluble entre les données et les annotations, donnant ainsi moyen au chercheur de vérifier constamment la cohérence de son analyse et à un éventuel pair de vérifier l'interprétation des données. En revanche, pour les LS, l'absence d'une transcription impose le retour constant aux données brutes ou une confiance aveugle en la fiabilité des « gloses ». L'absence d'une ligne de transcription, quand elle est couplée à l'impossibilité d'accéder aux données brutes, peut empêcher la validation par les pairs des résultats scientifiques.

- 22 Un autre inconvénient majeur des « gloses » utilisées par les chercheurs de LS est leur incapacité de décrire la forme signifiante du signe, c'est-à-dire la manière dont les mouvements dans l'espace et les formes des différents articulateurs corporels produisent le sens. Par exemple, les études sur les LS montrent qu'une partie du lexique des LS serait « conceptuel », c'est-à-dire que la forme citationnelle de certains signes ne porterait aucune marque de différenciation catégorielle¹⁴ (Millet, 2019) : c'est le cas du signe [BALAI/BALAYER] en LSF, dont la forme et le mouvement rappelle la manière d'empoigner le balai et de s'en servir¹⁵. Dans ces cas, la distinction entre noms et verbes ne s'actualise que dans le discours, et elle tient aux caractéristiques du mouvement des différents articulateurs (avec un mouvement plus large lorsque le signe signifie l'action de balayer ; Millet, 2006). Une fois qu'il a repéré cette différence de mouvement, le chercheur peut attribuer plus exactement l'étiquette [BALAI] ou [BALAYER], mais il n'empêche que ces « gloses » gommant les caractéristiques formelles du signe qui ont permis de déterminer cette attribution. Il devient alors difficile de repérer les régularités structurelles des LS pouvant expliquer la distinction entre nom et verbe. Seule l'association des gloses à une transcription détaillée et fidèle de la forme des signes permettrait de repérer les relations forme-sens en LS.
- 23 Reconnaissant l'importance de certains éléments formels, les spécialistes de LS agrémentent souvent les gloses de certaines indications pertinentes (2d') sur la forme des signes ; ils peuvent rendre plus évidente la multilinéarité de la LS (2d'') en réalisant une annotation sur partition. Mais dans les deux cas, les données représentées sont très partielles, bien plus que ne pourrait l'être une transcription en API d'une production en LV¹⁶, puisque seules certaines formes sont prises en compte (en se limitant souvent aux mains). Ainsi, dans ce script, (2d') et (2d'') donnent des indications sur la forme des mains (« proforme de forme S ») et indiquent des changements d'emplacement (« A>C » et « B>C »), mais ne donnent aucune indication sur les caractéristiques du mouvement réalisé. Or, en LS, la trajectoire ou la vitesse du mouvement caractérisent la signification du signe, et permettent de distinguer, dans l'exemple proposé dans le script de la Tab.2, la manière dont la poursuite se déroule.
- 24 Enfin, même les étiquettes choisies peuvent être source de confusion sur les informations codées. Par exemple, en signant « le chat poursuivait le chien », le signeur introduit de la main gauche le signe signifiant chien (dont on ignore la forme, car non indiquée), puis il réalise une proforme¹⁷ qui est identifiée comme étant « de forme S ». Un lecteur non averti pourrait imaginer l'existence d'une convention bien établie

faisant en sorte que toute configuration ayant cette forme soit toujours labélisée « S ». Toutefois, cette forme (que le lecteur, encore une fois, ignore) correspond à la configuration utilisée pour représenter la lettre S dans l'alphabet dactylogique de la LIS, mais pas à celle utilisée en LSF¹⁸. Cette description peut donc être une nouvelle source de biais, à cause du code d'identification choisi.

- 25 Le recours systématique à des traductions en LV pour représenter non seulement le sens mais aussi la forme des signes est, encore aujourd'hui, une pratique largement admise dans les études sur les LS. Au cours de ses conférences, EAP comparait souvent cette utilisation acritique des « gloses » aux remarques faites par la foule admirant les habits neufs de l'empereur dans la fable d'Andersen : face à l'éloge unanime, bien que tout le monde remarque la nudité de l'empereur, il est difficile d'oser affirmer que celui-ci est nu ; de même, bien que les biais causés par les étiquettes verbales soient visibles à tous, peu de chercheurs semblent en contester ouvertement l'utilisation.

3. La critique des « gloses » à la lumière des progrès de la recherche linguistique

- 26 Depuis le décès de EAP, en 2011, de nombreuses innovations ont touché les Sciences du Langage et l'étude des LS. Parmi ces nouveautés, le développement – récent – des principes F.A.I.R. (Findable, Accessible, Interoperable, Reusable) de la Science Ouverte, la diffusion de logiciels d'annotation électronique par partition, de plus en plus performants et maniables, et la naissance de nouvelles technologies d'enregistrement des données.
- 27 Depuis la diffusion du document « FAIR guiding principles for scientific data management and stewardship » (Wilkinson *et al.*, 2016), différentes instances nationales et internationales ont œuvré pour garantir une plus grande transparence et accessibilité des résultats de la recherche scientifique, qu'il s'agisse des sciences humaines et sociales ou des sciences dites « dures ». L'Union Européenne, à travers la « Amsterdam Call for Action on Open Science » de 2016, reprise en France par le « Plan National pour la Science Ouverte » de 2018, définit deux objectifs principaux : « permettre l'accès libre et total, d'ici 2020, à toutes les publications scientifiques » et « rendre systématique la possibilité de partage et de réutilisation des données pour toutes les recherches financées par des fonds publics » (MESRI-CoSO, 2018). La conséquence est la création de plateformes de dépôt d'articles scientifiques accessibles gratuitement par tous (par exemple les Archives Ouvertes HAL¹⁹), et la création de nombreux espaces numériques où les chercheurs peuvent déposer leurs corpus, logiciels, algorithmes, grilles d'analyse etc. pour les rendre accessibles aux autres chercheurs français et étrangers (mais non nécessairement à toute la société civile). Toutefois, la Science Ouverte n'impose pas l'ouverture sans conditions de toutes les données : les publications et les données doivent être « aussi ouvertes que possible, fermées autant que nécessaire » (CNRS, 2020). Il est donc, par exemple, plus important d'assurer la non-diffusion de données personnelles et/ou sensibles (en obtempérant au Règlement UE 2016/679 [Union Européenne, 2016] – plus connus sous le sigle de RGPD) plutôt que de rendre accessible les vidéos des corpus. Afin de garantir les droits des locuteurs ne désirant pas être identifiés, les corpus de LV peuvent subir un minutieux travail d'anonymisation avant d'être rendus accessibles : visages floutés, informations personnelles bippées, parfois, même, voix déformées. Mais pour les LS masquer le visage

des locuteurs équivaut à détruire une partie fondamentale et foncièrement linguistique des données. Dans les corpus de LS il est donc, en général, plus complexe de réunir les conditions nécessaires à une libre diffusion des données, qui ne peut avoir lieu que si les locuteurs donnent leur consentement éclairé à la diffusion de leur image reconnaissable. Enfin, la demande d'appliquer systématiquement les principes F.A.I.R. est relativement nouvelle dans le panorama des sciences du langage et, comme toute nouvelle manière de procéder, elle nécessite l'établissement de bonnes pratiques et la prise en main d'instruments qui, souvent, ne font pas encore partie de la formation scientifique des chercheurs. Ainsi, malgré les bonnes intentions de la plupart des personnes et des organismes impliqués, la Science Ouverte ne garantit pas encore la pleine accessibilité à tous les corpus de LS (et de données multimodales en général).

- 28 L'application de la Science Ouverte permet aux chercheurs de travailler sur des corpus produits, annotés et déposés en ligne par d'autres chercheurs. En imaginant une situation idéale où le corpus et ses annotations serait pleinement accessible, il est possible de revenir à l'exemple « le chat poursuivait le chien » (§2). Ayant à disposition les données brutes déposées, le chercheur travaillant sur le Romanesco et celui travaillant sur la LIS romaine pourront remplacer la ligne « source » (Tab.1a et Tab.2a) par un lien vers le corpus en ligne. Si toutefois, à cause par exemple d'une mauvaise connexion ou simplement du manque de temps, les chercheurs ne peuvent momentanément pas accéder aux données brutes, ceux travaillant sur la LV auront tout de même accès à une représentation assez fidèle des données brutes sur la ligne « transcription » (Tab.1b), tandis que ceux qui travaillent sur la LIS ne bénéficieront d'aucun lien vers les données. Cela veut aussi dire que si la pérennisation du corpus de LS n'est pas assurée, il n'existera dans 10-15 ans plus aucune trace de la forme signifiante de l'exemple en question, alors qu'il restera toujours une trace écrite et suffisamment fidèle du contenu du corpus de LV orale. La diffusion des corpus est donc une façon technologiquement avancée de pallier les difficultés liées à la représentation graphique, sans toutefois apporter la solution à ce problème.
- 29 Une autre amélioration substantielle ayant eu lieu ces 10 dernières années est liée aux logiciels d'annotation électronique par partition : instruments réservés, à l'époque de EAP, à une petite élite de chercheurs férus de nouvelles technologies, ils sont désormais de plus en plus utilisés par les experts de LS et par ceux qui, plus généralement, s'intéressent à la multimodalité. Qu'il s'agisse de ELAN (2002) ou d'un autre logiciel, ils permettent de créer une partition sur laquelle chaque phénomène peut être inscrit dans une ligne distincte et chaque étiquette d'annotation peut être reliée à un instant précis de la vidéo ou à une autre étiquette. Toutes les données, écrites en LV, sont requêtables soit directement sur le logiciel soit en exportant les annotations vers un tableur comme Excel ou un logiciel de traitement de données linguistiques. La richesse des annotations, la possibilité de regarder simultanément différents niveaux d'analyse et le lien permanent avec les données brutes, qui sont offerts par ces logiciels, permettent au chercheur de remarquer des relations entre phénomènes linguistiques qui, sans cela, passeraient probablement inaperçues.
- 30 Toutefois, bien que très sophistiqués, ces logiciels ne font que reproduire des étiquettes, qui peuvent être connectées et complexes, qui sont couplées à la vidéo, mais qui n'en restent pas moins des étiquettes verbales. ELAN permet donc de rechercher des étiquettes en associant celles-ci aux bons segments de vidéo, mais — par exemple — ne permet pas au chercheur d'avoir le même « coup d'œil » sur le contexte de

réalisation d'un signe que celui qu'il aurait en interrogeant un corpus de LV. En effet, en imaginant une procédure simple comme l'observation du contexte gauche (ou droit) d'un signe, le chercheur doit physiquement faire défiler le segment de vidéo correspondant à chaque contexte (Fig. 1), alors que dans un corpus de LV il pourra voir d'emblée une multitude d'occurrences avec leurs contextes (Fig. 2). Si, par chance, la transcription de la forme des signes a été réalisée, il doit tout de même lire des dizaines d'étiquettes décrivant chaque paramètre, là aussi, avec une dépense de temps et d'énergie cognitive bien supérieure à celle nécessaire à la comparaison des contextes d'une forme en LV.

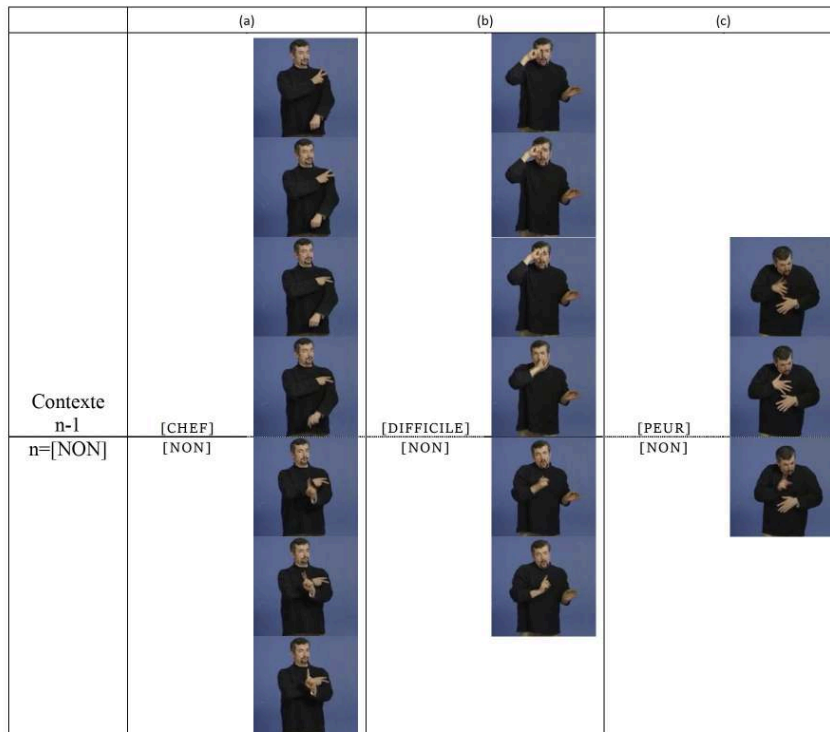


Fig. 1 – Contexte gauche (un signe) de trois occurrences du signe glosé comme [NON] extraits de la vidéo de « la recette de cuisine » de N. Chab (Corpus LS-Colin [Cuxac et Boutet, 2014]) : en l'absence d'un système de représentation graphique, la seule solution est l'observation de la vidéo (ici représentée par une capture d'écran toutes les 250 ms). [Étiquetage réalisé par Sallandre (2003) et réaffecté au bon segment de vidéo par nous]

	Contexte n-3	n=NON	Contexte n+3
(a)	// Médecins	NON	agréés, factures pour ...
(b)	... factures pour procédures	NON	effectuées. //
(c)	// Offre	NON	annulable, non modifiable ...
(d)	// Offre non annulable,	NON	modifiable, non remboursable. //
(e)	... annulable, non modifiable,	NON	remboursable. //
(f)	... également du bois	NON	broyé qui est ...
(g)	... exempt de décolorations	NON	fongiques et qui ...
(h)	... des cellules de parenchyme	NON	viables. //
(i)	// Ces solutions sont	NON	toxiques, non irritantes, ...
(j)	... sont non toxiques,	NON	irritantes, isotoniques, possèdent ...
(k)	... à spectre large	NON	spécifique et un ...

Fig. 2 – Contexte droit et gauche (3 mots) du mot [NON] dans un corpus éphémère constitué d'une sélection de résultats à la requête « non » dans le dictionnaire contextuel en ligne Reverso Context®

- 31 Ces dernières années, une nouvelle technologie d'enregistrement des données gestuelles commence à se diffuser : la capture automatique de mouvement (MoCap), dont les coûts ont considérablement baissé tandis qu'elle gagnait en maniabilité, devenant ainsi un instrument accessible aux linguistes. Avec ces technologies²⁰, la position des différents segments corporels mis en mouvement lors du déploiement des signes est enregistrée et peut être reversée dans ELAN ou dans d'autres logiciels d'annotation. Toutefois, la MoCap fournit ses données sous forme de chiffres (au minimum les coordonnées cartésiennes de chaque capteur pour chaque unité de temps, ce qui génère rapidement plusieurs centaines de milliers d'étiquettes), qui peuvent parfois être visualisés sous forme de courbes. Les données issues de la MoCap sont donc plus similaires à des « diagrammes de forme d'onde » du mouvement qu'à des transcriptions de celui-ci ; il ne s'agit donc pas, encore une fois, d'une solution au problème soulevé par EAP, puisque la forme des signes est « capturée » mais non représentée par les technologies de MoCap.
- 32 Bien que l'application des principes F.A.I.R. améliore l'accessibilité aux données brutes par les pairs, et que la diffusion des nouveaux instruments d'annotation et de saisie semblent révolutionner la manière de faire de l'analyse linguistique des LS et de la gestualité, aucun d'eux ne répond aux critiques soulevées par EAP sur l'utilisation des étiquettes verbales : seul un système de représentation graphique des LS peut fournir une solution à ce problème.

4. Les solutions graphiques pour représenter les LS à l'époque de EAP

- 33 Même si elle en a fait une de ses principales batailles, EAP n'est pas la seule linguiste à avoir souligné le problème de l'absence d'un système graphique permettant de transcrire les LS. Chez les chercheurs qui s'y sont intéressés, on recense principalement trois types d'approche : une approche théorique (fournissant une liste des caractéristiques que devrait posséder un système graphique pour remplir les fonctions d'un système d'écriture et/ou de transcription), une approche pratico-adaptative (testant un système existant – par exemple l'un des systèmes nés pour écrire les LS – pour vérifier ses performances comme système de transcription) et, enfin, une approche pratico-inventive (définissant un nouveau système de transcription permettant, selon son auteur, de satisfaire les exigences de représentation des LS).
- 34 Les solutions pratico-inventives qui existaient déjà à l'époque de EAP (parmi lesquelles SignFont [Newkirk, 1989], HamNoSys [Prillwitz *et al.*, 1989] et ASL-phabet [Supalla et Blackburn, 2003]; Fig. 3) sont toutes plus ou moins inspirées de la notation linéaire et paramétrique proposée par Stokoe (1960). Les différences entre les systèmes résident principalement dans le choix des paramètres représentés, le nombre de caractère du système et la forme graphique donnée aux caractères, ce qui influence respectivement la capacité du système à reproduire des signes en discours, sa finesse descriptive et enfin sa lisibilité et scriptibilité.

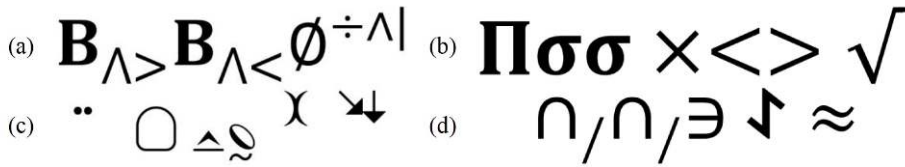


Fig. 3 – Exemples de solutions pratico-inventives pour représenter le signe ASL [MAISON]²¹ : (a) Notation de Stokoe, (b) SignFont, (c) HamNoSys et (d) ASL-phabet.

35 Parmi les systèmes issus de l’approche pratico-inventive, HamNoSys (Prillwitz *et al.*, 1989) est le seul à avoir été, et à demeurer encore aujourd’hui, largement utilisé pour la transcription de corpus en LS, bien qu’il ait été développé à l’origine pour des études lexicographiques. Composé d’environ 200 symboles, sa force réside dans son intégration informatique, qui assure la requêtabilité des informations qu’il code. Toutefois, la limite de cette requêtabilité se situe en amont du codage : si une forme n’est pas prévue par le système, elle ne peut être représentée, limitant de fait la finesse descriptive de HamNoSys et son adaptabilité à d’autres LS ou d’autres formes de gestuelle. À cela s’ajoute une autre faiblesse, commune à tous les systèmes descendants de la Notation de Stokoe : la disposition linéaire de blocs paramétriques codés indépendamment l’un de l’autre – même si associés à des symboles relativement iconiques comme dans le cas de HamNoSys – semble empêcher le chercheur de relire aisément leurs transcriptions. L’explication de ce phénomène pourrait résider dans le contraste entre la production linguistique multilinéaire et simultanée des différents paramètres et leur reproduction graphique monolinéaire et séquentielle. Dans certains cas, cette segmentation et cette linéarisation sont même internes aux paramètres : le mouvement, par exemple, est souvent découpé en sous-paramètres tels que la forme, la direction, la dynamique, la répétition, le contact, etc. L’absence d’une vision synthétique des paramètres, voire de tout le signe, rend la relecture très similaire à un long processus de déchiffrement, empêchant le « coup d’œil » sur les données de LS que le chercheur s’intéressant aux LV peut – quant à lui – avoir (comparer Fig. 1 et Fig. 4).

	(a)	(b)	(c)
Contexte n-1	[CHEF] V _T V ^{T*} ∅∅ A _T ∅∅	[DIFFICILE] G _V ∧*∅∅ F _∅ ∅∅	[PEUR] 5 _T []IX [∅] ∅ 5 _T []IX [∅] ∅
n=[NON]	[NON] V _T V ^{∅*} ∅∅ G _∅ ∅∅	[NON] G _{<} ∅∅ V _∅ ∅∅ F _∅ ∅∅ G _{<} ∅∅	[NON] 5 _T []∅X

Fig. 4 – Transcription en Notation de Stokoe (en version adaptée par Volterra, 1987) des extraits de la Fig. 1 :

Contexte gauche (un signe) de trois occurrences du signe glosé comme [NON] extraits de la vidéo de « la recette de cuisine » de N. Chab (Corpus LS-Colin [Cuxac et Boutet, 2014]).

[Transcription en notation de Stokoe par nous ; la verticalité permet de mieux prendre en compte la latéralité]

36 Bien que ces systèmes fassent notablement progresser la recherche d’une solution au problème de la représentation des LS, ils ne permettent toujours pas d’avoir un système complet, lisible et facilement exploitable. Dans un article de 2010, EAP et ses collaborateurs résument ainsi les critiques portées à l’ensemble des systèmes issus de l’approche pratico-inventive, dont ils avaient étudié les caractéristiques :

« Grande complexité dans l’écriture mais aussi dans la lecture ; manque de connexion avec la forme effective du signe ; grande difficulté, voir impossibilité, de transposer des discours en signes ; omission de parties importantes des signes, telle que les composantes non manuelles ; difficultés de mémorisation des critères

utilisés pour transposer les signes ; quasi impossibilité de réutiliser les données transcrites dans des buts différents de ceux pour lesquels la transcription a été effectuée ; difficultés à rendre la dynamique des signes et l'utilisation particulière de l'espace de signation. » (Bianchini et al., 2011 ; p. 72-73)

- 37 En alternative, EAP propose une approche pratico-adaptative au problème de la représentation graphique des LS, expérimentant l'utilisation de SignWriting²² (Sutton, 1995), un système graphique développé par la chorégraphe Valerie Sutton à partir de 1974. Sutton définit son système comme une écriture (et non une transcription) des LS, dont le but est de permettre aux sourds de mettre noir sur blanc leurs productions langagières, comme ils pourraient le faire avec une LV écrite. EAP propose de tester son utilisation tant comme système d'écriture que de transcription des LS.
- 38 Du point de vue organisationnel, SW est composé d'un ensemble de symboles (appelés par EAP et son équipe « glyphes ») qui sont : (1) standardisés ; (2) fortement iconiques²³ ; (3) agencés dans un espace bidimensionnel (qu'ils appellent « vignette ») qui représente analogiquement l'espace de signation (Boutet et Garcia, 2007) ; (4) capables de représenter tant les composantes manuelles (forme, orientation et mouvement des mains) que non manuelles (balancements du corps, direction du regard, expressions faciales) des signes. Le plus grand avantage de SW est sa lisibilité : si le travail d'écriture nécessaire à créer une vignette en SW est assez laborieux, la lecture/relecture des textes est, quant à elle, très rapide et aisée (Gianfreda *et al.*, 2009), car SW offre une vision synthétique de chaque paramètre et de l'ensemble du signe, qui respecte la simultanéité et la multilinéarité des LS.
- 39 Les études menées par EAP et son équipe – qui sont résumées dans la thèse de Bianchini (2012) – ont souligné divers avantages liés à l'utilisation de SW comme système d'écriture des LS. Grâce à sa lisibilité, SW permet de mener des analyses linguistiques qualitatives, offrant au linguiste le « coup d'œil » sur les données dont il était question plus haut (comparer Fig. 1, Fig. 4 et Fig. 5). SW a ainsi fait émerger, pour la première fois, de nombreuses réflexions métalinguistiques sur les LS (Bianchini, 2021), réflexions qui n'auraient pas été possibles sans un système graphique offrant une vision synthétique et lisible des signes.

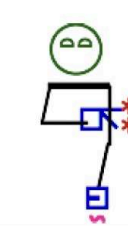
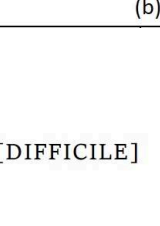
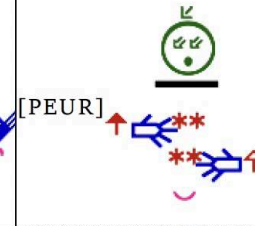
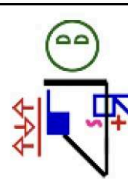
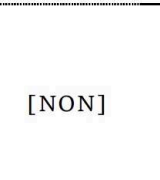
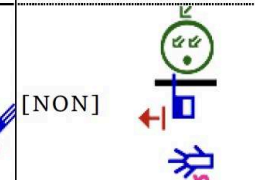
	(a)	(b)	(c)
Contexte n-1	[CHEF] 	[DIFFICILE] 	[PEUR] 
n=[NON]	[NON] 	[NON] 	[NON] 

Fig. 5 – Transcription en SignWriting des extraits de la Fig. 1 : Contexte gauche (un signe) de trois occurrences du signe glossé comme [NON] extraits de la vidéo de « la recette de cuisine » de N. Chab (Corpus LS-Colin [Cuxac et Boutet, 2014]). [Transcription réalisée par nous avec SignMaker (Sutton, 2010) ; coloration proposée par SignMaker] [La disposition verticale des vignettes permet de mieux prendre en compte la latéralité des signes]

- 40 Les tentatives d'utilisation de SW comme moyen de transcription se heurtent toutefois aux exigences de la linguistique de corpus en matière d'analyse quantitative. En effet, s'il est possible de transcrire avec SW un corpus de LS, celui-ci sera difficilement exploitable. Le problème principal réside dans l'absence (du moins pour l'instant) de règles de composition « orthographiques ». Doté d'environ 35 000 glyphes pouvant être déposés quasi-librement dans l'espace graphique (ce qui constitue une force en termes de lisibilité), SW propose au scripteur une infinité de solutions différentes pour coder un même signe : la différence entre un signe bien écrit et un signe à corriger tient alors uniquement à sa lisibilité (Bianchini *et al.*, 2021), c'est-à-dire que si le signe est lu de façon conforme à l'intention du scripteur, quelle que soit la combinaison des glyphes choisis et leur disposition dans l'espace graphique, il sera jugé correct. Aussi est-il non seulement extrêmement complexe de parvenir à un accord inter-annotateurs satisfaisant, mais aussi impossible de requêter informatiquement toutes les occurrences d'une forme (à moins de connaître toutes les combinaisons permettant de la coder, ce qui est techniquement impossible). Afin de pouvoir recenser tous les signes ayant une certaine forme, la seule solution semble être de leur associer une étiquette verbale qui identifie en LV les signes ayant une même caractéristique formelle²⁴, mais cela revient à utiliser les « gloses », qu'EAP combattait.
- 41 Un autre problème posé par SW dans son utilisation en tant que système de transcription concerne son informatisation²⁵. En premier lieu, l'utilisation de SignMaker (Sutton, 2010), le logiciel d'édition officiel de SW, est très laborieuse : alors qu'écrire un signe à la main peut prendre une ou deux minutes, ce temps est multiplié par trois ou quatre lorsque le signe doit être écrit informatiquement. Par ailleurs, bien que SW soit le seul système de représentation graphique des LS ayant été reconnu par le Consortium Unicode – ce qui lui permettrait de fonctionner sur n'importe quel logiciel – cette reconnaissance n'a pas été assortie de la création d'une police de caractères permettant de maintenir la disposition des glyphes en vignette²⁶, élément clé de la lisibilité de SW. Concrètement, à ce jour, SW ne peut donc être utilisé dans un logiciel d'annotation tel que ELAN, ce qui rend impossible son exploitation pour analyser des corpus.
- 42 En l'état actuel des choses, bien que son utilisation soit problématique, SW est néanmoins le seul système qui semble pouvoir aspirer à devenir un système d'écriture pour les LS : la preuve en est que les tentatives plus récentes de créer une écriture des LS s'inspirent largement du fonctionnement de SW (c'est le cas, par exemple, de Si5s [Augustus *et al.*, 2014]). On peut donc émettre l'hypothèse que, dans un avenir plus ou moins lointain, SW (ou l'un de ces dérivés) pourra être accepté par la communauté sourde, ce qui permettra de développer des règles d'« orthographe » plus strictes et poussera vers une amélioration de son informatisation (Bianchini, 2013). Cela pourrait à terme (mais cela reste hypothétique) permettre de l'utiliser aussi comme moyen de transcription ou comme base pour la création d'un système de transcription performant pour les LS.

5. Critères de développement d'un système de transcription des LS

43 En croisant les avantages et inconvénients de divers systèmes graphiques utilisés pour les LS (dont une partie ont été présentés ci-dessus), il est possible d'établir une liste (non exhaustive) de caractéristiques souhaitables pour un système de transcription²⁷ des LS (et de la gestualité co-verbale). Ces caractéristiques composent un système de représentation des LS qui véhicule la forme signifiante des signes tout en garantissant l'exploitation des corpus transcrits avec ce système :

- Fidélité : le système doit reproduire de la manière la plus fidèle possible la forme signifiante des différents paramètres du signe mais aussi du signe dans son ensemble, afin de fournir une trace de la manière dont le signe a été réalisé ;

- Lisibilité : l'utilisateur doit être en mesure de relire aisément les transcriptions (réalisées par lui ou par d'autres) afin, entre autres, de pouvoir réaliser une analyse qualitative des données ;

- Scriptibilité : l'utilisateur doit être en mesure d'écrire (à la main et/ou au moyen d'instruments informatiques adéquats) le système de façon simple et rapide, voire automatique ou semi-automatique ;

- Requêtabilité : l'utilisateur doit pouvoir requêter facilement les informations qu'il a codées afin, entre autres, de pouvoir réaliser une analyse quantitative des données ;

- Rentabilité : la qualité et la quantité des informations codées, et donc requêtables, doit compenser l'utilisateur du temps investi à transcrire les données²⁸.

44 En plus de ces caractéristiques, le système doit pouvoir s'adapter à l'ensemble des utilisateurs potentiels. Il est alors possible d'ajouter les éléments suivants :

- Universalité : l'utilisation du système ne doit pas être liée à la connaissance d'une LS ou d'une LV particulière de la part du chercheur ;

- Expansibilité : le système doit pouvoir être utilisé – sans besoin de refontes – pour représenter toutes les LS (et pas seulement celles pour lesquelles il a été initialement prévu), voire pour représenter d'autres pratiques langagières visuo-gestuelles (par exemple la gestualité co-verbale) ;

- Indépendance : le système ne doit pas être lié à un objectif de recherche ou à un paradigme spécifique, afin de pouvoir être utile à des chercheurs suivant différentes approches et cadres théoriques ;

- Multiplicité : le système ne doit pas être conçu pour n'être utilisé que par des chercheurs s'intéressant à une branche spécifique de la linguistique (notamment ceux spécialisés en phonologie des LS, qui sont normalement à l'origine des systèmes graphiques) mais doit permettre d'accéder aux données selon le niveau d'analyse souhaité.

45 Enfin, le système de représentation graphique doit s'adapter à différentes contraintes externes, surtout celles liées aux progrès technologiques. La création d'un système de transcription est un travail de longue haleine : ne pas prendre en compte les innovations technologiques rendrait obsolète le système avant même sa mise en œuvre. La liste peut donc être complétée par les points suivants :

- Actualité : le système doit être conçu en tenant compte de la croissante importance des processus de standardisation des données numériques, en tirant avantage des possibilités offertes par la technologie actuelle et en prenant en compte les évolutions (des standards et des technologies) qui pointent déjà à l'horizon ; bien que l'actualité soit – par définition – une caractéristique éphémère, être à la pointe de la technologie permet d'assurer une plus grande longévité au système et facilite aussi son éventuelle mise à jour.

- Liberté : le système doit être disponible gratuitement et son utilisation/modification ne doit être soumise qu'à l'attribution de son origine intellectuelle et au maintien de cette liberté de diffusion (ce qui équivaut à lui attribuer une licence Common Creative CC BY-SA) ;

- Interopérabilité : le système doit fonctionner sur tous les logiciels (des éditeurs de textes et diaporamas aux logiciels d'annotation, en passant par les tableurs, les logiciels de dessin etc.), les navigateurs (Firefox, Safari, Chrome, etc.), les systèmes d'exploitation (PC, Mac, Linux, Android, etc.) et les dispositifs (téléphone, tablette, ordinateur, etc.), présents et— si possible —futurs ;

- Pérennité : le système doit se baser sur des pratiques permettant d'assurer tant la pérennité du système même (au-delà de la durée de vie de l'équipe de recherche l'ayant créée et des technologies utilisées pour le développer) que des données ayant été produites avec celui-ci, quelles que soient les percées de la technologie dans les années à venir.

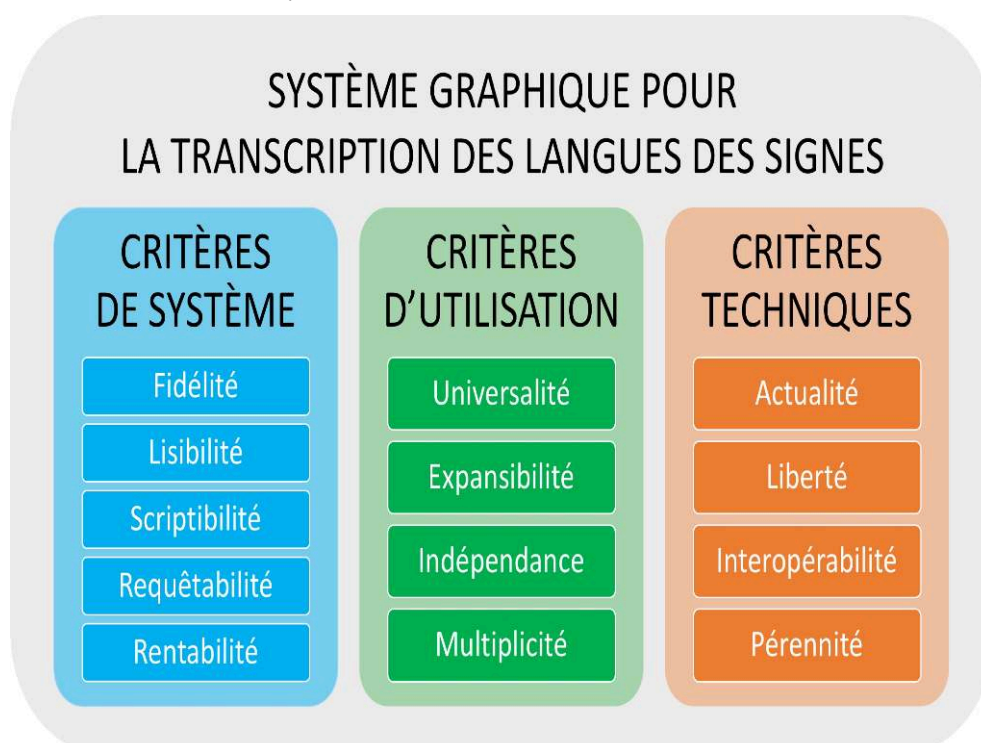



Fig. 6 – Critères pour la réalisation d'un système graphique adapté à la transcription et à l'exploitation de corpus de LS (et de gestualité co-verbale)

- 46 Concentrer dans un seul système toutes ces caractéristiques (Fig. 6) nécessite de repenser la manière d'approcher la construction d'un système graphique des LS. Il faut adopter une approche « grapholinguistique » où le point de vue linguistique (concernant les informations que les linguistes veulent pouvoir coder et ensuite requêter) est toujours mis en relation avec l'aspect graphique (touchant la manière dont les systèmes graphiques permettent de coder ces informations). Un exemple de système développé dans une approche grapholinguistique est Typannot (Boutet *et al.*, 2018 ; Doan *et al.*, 2019), qui permet de transcrire les LS et la gestualité co-verbale. Il voit le jour au sein de l'équipe de recherche GestualScript, établie au sein de l'École Supérieure d'Art et de Design d'Amiens.

6. Nouvelle proposition pour la transcription des LS : Typannot

- 47 L'équipe GestualScript vise la création d'un système de notation des LS et de la gestualité qui assure la description phonologique des paramètres des LS tout en garantissant la *fidélité*, la *lisibilité*, la *scriptibilité*, la *requêtabilité*, la *rentabilité* et l'*expansibilité*²⁹ du système. Alors que les équipes suivant une approche pratico-inventive sont généralement composées de linguistes travaillant avec l'appui d'informaticiens, l'équipe GestualScript identifie dans le savoir-faire professionnel des graphistes et des dessinateurs de caractères typographiques un apport fondamental à la constitution d'un système graphique utilisable et efficace. Dans cette équipe, graphistes, typographes et informaticiens ne sont plus des spécialistes qui répondent aux cahiers des charges élaborés par les linguistes, mais des membres actifs à toutes les étapes menant à la définition d'un système graphique de description phonologique des LS et de la gestualité.
- 48 Le résultat de cette approche grapholinguistique est un système *typographique* de transcription, c'est-à-dire une *famille de polices de caractères* permettant la représentation graphique des LS et de la gestualité. S'inscrivant dans l'univers de la typographie numérique, Typannot base son fonctionnement sur une distinction fondamentale de ce domaine, celle entre les caractères typographiques et les glyphes. Un *caractère typographique* est une information, une unité abstraite, qui peut être (si le système d'écriture est suffisamment stable et considéré utile à la communauté) identifié par un code attribué par le Consortium Unicode³⁰; ainsi en appuyant simultanément les touches {SHIFT} et {A} du clavier, c'est le caractère identifié par U+0042 (labellisé LATIN CAPITAL LETTER A) qui est renseigné dans le logiciel d'édition, et c'est aussi U+0042 qui est requêté lorsqu'on demande à un logiciel de traitement de texte de restituer tous les mots comprenant un A majuscule. Ce caractère abstrait ne peut être affiché à l'écran que s'il est associé à un *glyphe*, une unité graphique concrète dessinée au sein d'une police de caractère³¹. Ainsi U+0042 pourra faire apparaître un A, A ou A³² selon la police installée. Un glyphe peut aussi représenter l'union de plusieurs caractères, on parlera alors de *ligature typographique*. Taper la touche {O} suivie de {E} équivaut à insérer les caractères U+006F et U+0065; si cette opération a lieu en utilisant une police de caractères supportant les ligatures, alors ce ne seront pas les glyphes [o] et [e] qui apparaîtront, mais la ligature [œ] (Tab.3a). Les ligatures ne concernent pas exclusivement les lettres et une ligature peut être graphiquement très différente de la somme des glyphes qui la composent : par exemple les caractères U+1F469, U+1F3FB, U+200D et U+1F3A4 correspondent respectivement aux glyphes représentant une femme, la couleur rose pale, un espace 'sans largeur'³³ et un micro, mais s'ils sont écrits l'un à la suite de l'autre, c'est la ligature représentant une chanteuse au teint pâle qui apparaît (Tab.3b). L'existence des ligatures exerce une grande influence sur la requêtabilité des informations présentes dans le texte : formuler une requête textuelle demandant d'extraire toutes les occurrences de U+006F restitue tant les mots contenant le glyphe [o] que ceux contenant [œ]; de même requêter U+1F469 restituera tant le glyphe de la femme que celui la chanteuse.

	(a)	(b)
--	-----	-----

Numéros Unicode des caractères insérés pour former une ligature	U+006F	U+0065	U+1F469	U+1F3FB	U+200D	U+1F3A4
Glyphes correspondants à chaque caractère utilisé	[o]	[e]	[♀]		[]	[♀]
Ligature obtenue à partir des caractères insérés	[œ]		[♀]			

Tab.3 – Une combinaison de caractères Unicode peut être affichée sous la forme d'une suite de glyphes ou d'une ligature typographique

- 49 Typannot est tout d'abord une liste d'informations — chacune étant inscrite dans un caractère typographique appelé *caractère générique* — qui permet de décrire HS, MAct, EAct, LOCini et Mov. Chaque paramètre est décomposé en sous-parties dont les formes sont décrites le plus objectivement possible, à travers l'identification de traits articulatoires squelettiques et/ou musculaires. Par exemple, la configuration de la main est décrite à travers la forme et l'angle de chaque doigt et du pouce, l'interaction des doigts entre eux et avec le pouce, et la position du pouce par rapport à la paume. Ainsi constitués, les caractères génériques permettent d'assurer la *fidélité*, l'*universalité*, l'*expansibilité* et l'*indépendance* de Typannot.
- 50 Traiter ces génériques non seulement comme informations linguistiques mais aussi comme caractères typographiques permet à Typannot de profiter de la politique de standardisation mise en œuvre par le Consortium Unicode, l'entité internationale en charge de garantir l'interopérabilité et la pérennité de tous les textes numériques, quel que soit le système d'écriture avec lequel ils ont été écrits, en associant de manière indissoluble chaque caractère (et non glyphe) à un code standardisé. Bien que n'ayant pas encore reçu une reconnaissance officielle par le Consortium³⁴, Typannot est conçu pour répondre à l'ensemble des critères permettant cette admission et occupe, pour l'instant, une Zone à Usage Privée d'Unicode. Ceci permet aux caractères de Typannot d'être *interopérables* et *requêtables* à l'instar de tout autre élément du Standard Unicode, qu'il s'agisse d'un caractère de l'alphabet latin, d'un hiéroglyphe ou d'un autre caractère appartenant à un quelconque système d'écriture. Une fois Typannot reconnu par le Consortium, la *pérennité* du système sera garantie, car toute évolution d'Unicode sera appliquée par le Consortium à tous les systèmes graphiques (y compris Typannot) dont il reconnaît l'existence. Enfin, concevoir Typannot en termes de caractères typographiques en fait un système foncièrement libre de droits, car personne ne peut posséder des droits sur l'alphabet latin, le Hangul ou les hiéroglyphes.

- 51 Pour les exigences des linguistes, la liste des traits articulatoires — inscrits dans les caractères génériques — qui décrivent les divers paramètres pourrait sembler suffisante ; toutefois, les caractères génériques ne peuvent être affichés à l'écran que s'ils sont associés à des glyphes. Les *glyphes génériques* (Fig. 7) permettent donc de rendre visibles les caractères génériques : leur forme graphique épuré et fortement distinctive en assure la *lisibilité* et la *scriptibilité* ; le choix de représenter schématiquement le contenu informationnel, sans liens avec une LS ou une LV spécifique, garanti l'*universalité* de Typannot.

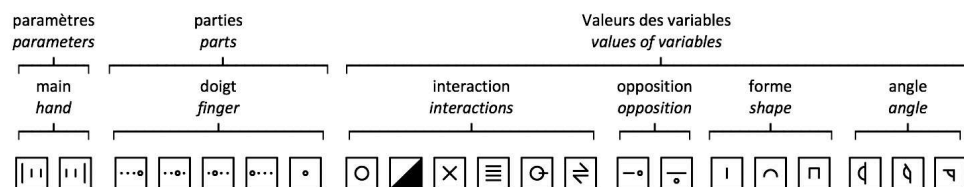


Fig. 7 – Inventaire des 21 caractères génériques de Typannot HandShape

- 52 Pour représenter une occurrence d'une configuration de la main, une sélection de caractères (et donc les glyphes) génériques est insérée dans une *formule graphématique* (Fig. 8) à la syntaxe rigide. L'idée de mettre les informations dans une formule existe depuis Stokoe, mais Typannot diffère de ses prédécesseurs par le niveau de détail de l'information codée et la possibilité de coder des formes imprévues. En effet, la notation de Stokoe offre une liste de configurations attestées, qui sont considérées comme des unités non décomposables et qui sont associées à un symbole. À l'inverse, Typannot code l'information au niveau infra-paramétrique et associe chaque partie de la main à ses caractéristiques morphologiques (forme, angle, interactions), n'offrant donc pas une liste préétablie de configurations attestées mais une liste de caractéristiques (21 au total) susceptibles d'être combinées pour obtenir toutes les configurations possibles. Ce choix permet une requêtabilité sur plusieurs niveaux : il est possible de rechercher une configuration de façon globale, mais aussi de rechercher — par exemple — toutes les configurations où deux doigts se croisent, indépendamment des doigts concernés, ou de la forme des autres doigts, ou de la position du pouce. Une configuration n'a donc pas besoin d'être attestée et prévue pour être transcrite, pas plus qu'il ne faut connaître toutes les configurations contenant un certain trait pour pouvoir le requêter. La structure de Typannot permet donc d'effectuer des analyses quantitatives tant au niveau du paramètre que de ses différentes sous-parties, augmentant les possibilités d'exploitation des corpus transcrits avec ce système et ce, à partir d'une liste extrêmement restreinte de caractères.

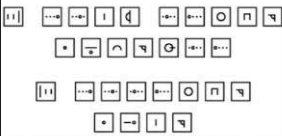
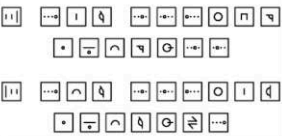
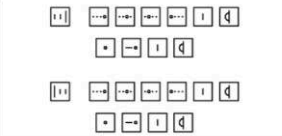
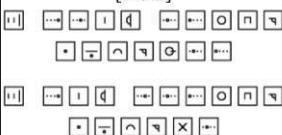
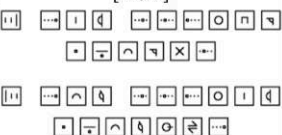
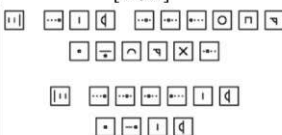
	(a)	(b)	(c)
Contexte n-1	[CHEF] 	[DIFFICILE] 	[PEUR] 
n=[NON]	[NON] 	[NON] 	[NON] 

Fig. 8 – Transcription en formule graphématique de Typannot HandShape des extraits de la Fig. 1 : Contexte gauche (un signe) de trois occurrences du signe glosé comme [NON] extraits de la vidéo de « la recette de cuisine » de N. Chab (Corpus LS-Colin [Cuxac et Boutet, 2014]).

[Transcription en caractères génériques par nous]

[NB : contrairement aux figures précédentes, seule la forme de la main est codée ici]

- 53 La possibilité de combiner des informations pour décrire toute configuration imaginable permet de limiter le nombre d'informations nécessaires à caractériser sa forme ; en contrepartie, la formule ne peut supporter une erreur de syntaxe sans compromettre la requêtabilité des informations. Pour cette raison, l'équipe GestualScript a mis au point la Typannot Keyboard (Fig. 9), un clavier virtuel qui permet une insertion rapide des informations et qui génère automatiquement la formule graphématique évitant ainsi toute erreur de syntaxe. Ce clavier permet aussi de vérifier la justesse de la transcription grâce à un avatar qui reproduit la formule insérée. Enfin, la Typannot Keyboard est conçue pour être couplée à des dispositifs de MoCap, ce qui permettra— à terme³⁵ —une transcription semi-automatique des corpus avec Typannot, améliorant la rentabilité du système.

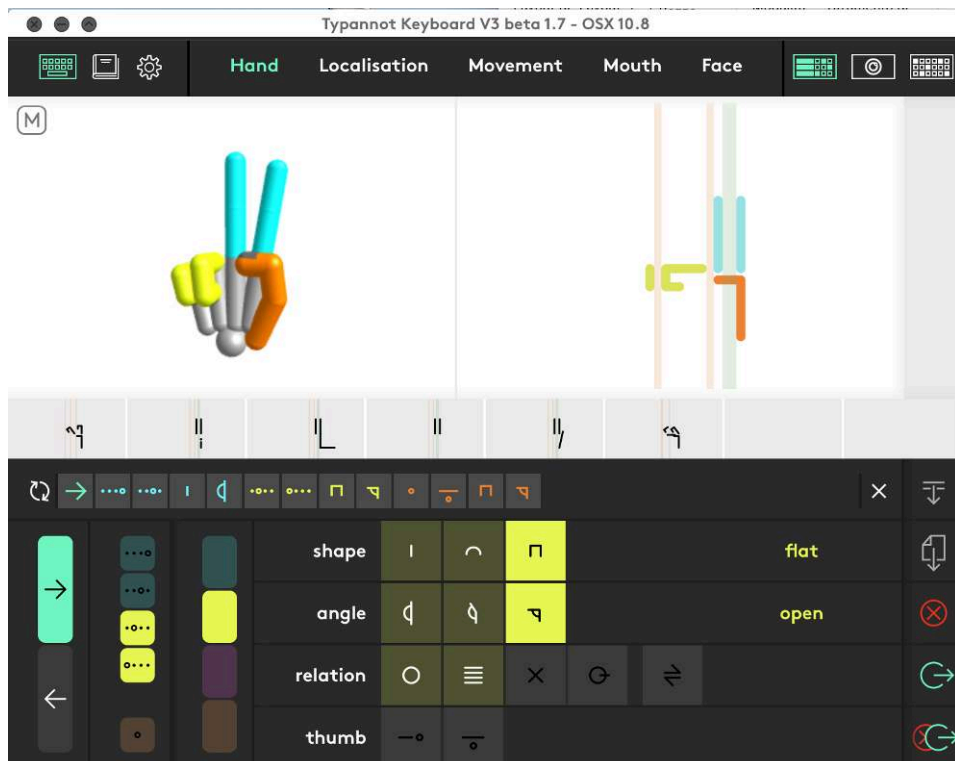


Fig. 9 – Interface d'insertion des configurations de la main dans Typannot Keyboard

54 Bien que les glyphes génériques soient lisibles et que la formule soit assez facile à déchiffrer, il manque la possibilité — offerte par SW — de pouvoir examiner les données de façon globale. Pour résoudre cette limite, l'équipe GestualScript fait appel aux ligatures typographiques. Chaque formule graphématique, composée de caractères génériques, peut être affichée sous deux formes distinctes mais co-existantes : la version étendue (Fig. 8) et la version composée (Fig. 10 et Fig. 11). Les *glyphes composés* sont le résultat de la ligature typographique des caractères génériques et offrent une vision synthétique et lisible du paramètre codé³⁶, ce qui garantit le « coup d'œil » nécessaire à l'analyse qualitative des données. Alors que le symbole de SignWriting est une unité informationnelle et graphique non ultérieurement décomposable, ici ce n'est que la visualisation qui est unitaire : chaque glyphe composé continue à contenir en son sein l'ensemble des informations des caractères dont il est la ligature, assurant donc une requêtabilité infra-paramétrique identique à celle de la version étendue. Le passage de la visualisation étendue à la visualisation composée et vice-versa n'est qu'une question d'affichage, réglé dans les paramètres des logiciels où la famille de polices de caractères Typannot est utilisée.






	(a)	(b)	(c)
Contexte n-1	[CHEF] 	[DIFFICILE]  i	[PEUR] 
n=[NON]	[NON] 	[NON] 	[NON] 

Fig. 10 – Transcription en glyphes composés de Typannot HandShape des extraits de la Fig. 1 : Contexte gauche (un signe) de trois occurrences du signe glosé comme [NON] extraits de la vidéo de « la recette de cuisine » de N. Chab (Corpus LS-Colin [Cuxac et Boutet, 2014]). [Pour une plus grande lisibilité, seulement les doigts saillants sont affichés ici] [Transcription en caractères génériques par nous, composition des glyphes par Morgane Rébulard] [NB : contrairement aux figures précédentes, seule la forme de la main est codée ici]

55 Le système étant en phase de développement, il n'est pas possible de fournir un exemple de ce que donnerait un signe entier codé avec Typannot. Une fois que le système sera terminé, la double visualisation de Typannot assurera la *multiplicité* du système, car elle permettra aux chercheurs s'intéressant à différents niveaux d'analyse, de la phonétique à la pragmatique, d'accéder aux données dans le format le mieux adapté à leurs exigences. Il serait également possible, sans modifier le contenu informationnel des caractères génériques, de proposer une autre forme de visualisation plus adaptée aux besoins d'un groupe de recherche (par exemple, des chercheurs travaillant sur la phonologie des LS pourraient désirer un affichage de tous les doigts, Fig. 11), ou plus en adéquation avec une technologie aujourd'hui inconnue.







	(a)	(b)	(c)
Contexte n-1	[CHEF] 	[DIFFICILE] 	[PEUR] 
n=[NON]	[NON] 	[NON] 	[NON] 

Fig. 11 – Transcription en glyphes composés de Typannot HandShape des extraits de la Fig. 1 : Contexte gauche (un signe) de trois occurrences du signe glosé comme [NON] extraits de la vidéo de « la recette de cuisine » de N. Chab (Corpus LS-Colin [Cuxac et Boutet, 2014]). [Pour une plus grande précision, tous les doigts (saillants ou pas) sont affichés ici] [Transcription en caractères génériques par nous ; dessins des glyphes par Morgane Rébulard] [NB : contrairement aux figures précédentes, seule la forme de la main est codée ici]

Conclusion

- 56 En 2011 EAP a laissé en héritage des « *questions that need to be posed* ». Elle y pointait une lacune lourde de conséquences pour l'analyse des LS, à savoir l'absence d'une méthode efficace pour rendre compte de la forme signifiante des signes et le recours systématique à des étiquettes verbales pour l'éluder. Pour EAP, la solution résidait dans l'utilisation d'un système graphique permettant de figer de façon synthétique et lisible la forme signifiante des signes, comme le font les systèmes phonographiques pour les LV. Toutefois, il s'est avéré que le système qu'elle avait décidé de tester, SignWriting, très prometteur comme système d'écriture, n'était pas adapté à la transcription ni, surtout, à l'exploitation de vastes corpus de LS.
- 57 Aujourd'hui la situation ne semble guère avoir changé : le recours systématique aux étiquettes verbales est encore la pratique la plus répandue et il ne semble pas y avoir eu, à quelques exceptions près³⁷, une prise de conscience des biais que cette manière de faire impose à l'analyse des LS. Pour reprendre la métaphore de EAP sur « les habits neufs de l'empereur », la foule d'utilisateurs des « gloses » ne semble pas encore écouter les quelques chercheurs qui soulignent la nudité de l'empereur. Mais cette foule n'est pas composée que de chercheurs experts de LS. Si EAP ne s'intéressait qu'aux LS, il est important de noter que le problème qui la préoccupait concerne aussi les experts de gestualité, qui rencontrent les mêmes difficultés et pallient eux aussi l'absence d'un système de transcription efficace en se concentrant sur la fonction des gestes, plus que sur leur forme. Les questions de EAP ne sont donc pas seulement d'actualité pour les LS, mais aussi pour l'ensemble des chercheurs s'intéressant à la multimodalité. Les problèmes étant communs aux experts de gestualité et de LS, les solutions auraient grand intérêt à être, elles aussi, communes.
- 58 Toutefois, si pouvoir transcrire la partie gestuelle d'un corpus multimodal est bien un « problème de linguiste », il n'est pas dit que les linguistes puissent à eux seuls fournir une solution. Les tentatives axées uniquement sur la composante linguistique et informatique du problème ne semblent pas faire émerger de solutions réellement efficaces. En effet, la représentation graphique des LS est un problème graphique.

L'écriture n'est pas uniquement une forme dérivée du langage, c'est une technologie à part entière qui – bien qu'appliquée au langage – a ses règles, ses universaux, ses principes de fonctionnement qui vont bien au-delà des besoins du langage et qui doivent être pris en considération lors de la création d'un système graphique. La solution se trouve donc à la frontière entre la linguistique, l'informatique et la typographie, et elle attend de voir le jour. C'est à la charnière de ces disciplines que se situe le développement de Typannot, le premier système de transcription typographique des LS et de la gestualité co-verbale.

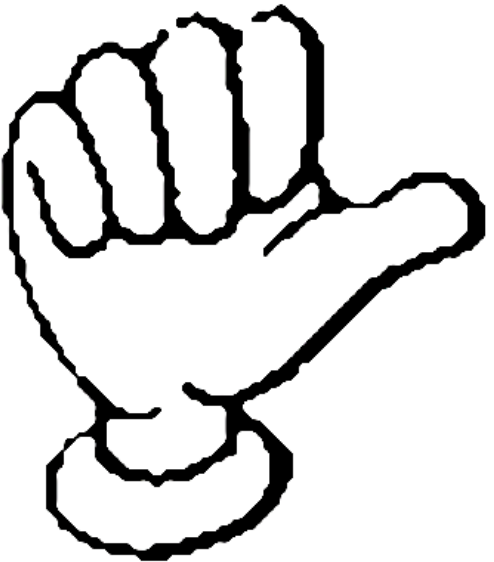

- 59 Le lecteur voudra bien m'excuser si, dans la conclusion de cette conclusion, je sors de la rigueur avec laquelle j'ai tenu à rédiger cet article, pour dire graphiquement merci à Elena de m'avoir initié à la recherche scientifique, et de m'avoir transmis son enthousiasme pour la quête d'une solution aux problèmes de la représentation graphique des LS.



« Merci beaucoup Elena » en LS italienne, écrite avec SignWriting

- 60 Note finale
- 61 Afin de satisfaire l'éventuelle curiosité du lecteur voici les formes correspondantes aux paramètres ou au signes entiers présentés dans l'article et qui n'ont pas pu être présentées le long du texte.

Description en LV	Configuration S en LIS	Configuration S en LSF
-------------------	------------------------	------------------------

Dessin de la forme		
Référence	§2, note 17	§2, note 17

- 62 NB : sur le dessin du signe [MAISON], les parties du corps grisées ou absentes ne sont pas représentées dans la transcription fournie dans la Fig.3.

BIBLIOGRAPHY

Antinoro Pizzuto E., I. Chiari & P. Rossini. 2008. The representation issue and its multifaceted aspects in constructing Sign Language corpora: questions, answers, further problems. In : O. Crasborn, E. Efthimiou, T. Hanke, E. Thoutenhoofd & I. Zwitterlood (eds.) *Proceedings of the 3th Workshop on Representation and Processing of Sign Languages, 6th International Conference on language resources and evaluation (LREC2008), Marrakech 01/06/2008*: 150-158.

Antinoro Pizzuto E. 2009. Meccanismi di coesione testuale e Strutture di Grande Iconicità nella Lingua dei Segni Italiana (LIS) e altre lingue dei segni. In : C. Bertone & A. Cardinaletti (eds.) « *Alcuni capitoli della grammatica della LIS* ». Cafoscarina (Venezia): 137-138.

Antinoro Pizzuto E., I. Chiari & P. Rossini. 2010. Representing signed languages: theoretical, methodological and practical issues. In : M. Pettorino, F.A. Giannini, I. Chiari & F. Dovetto (eds.) « *Spoken communication* ». Cambridge Scholars Publ. (Newcastle upon Tyne): 205-241.

Assemblée Nationale - Sénat. 2005. Loi n°2005-102 du 11 février 2005 pour l'égalité des droits et des chances, la participation et la citoyenneté des personnes handicapées.

[[www.legifrance.gouv.fr/affichTexte.do?](http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT00000809647&dateTexte=&categorieLien=id)

cidTexte = JORFTEXT00000809647&dateTexte = &categorieLien = id]

Augustus R.A., E. Ritchie & S. Stecker (E. Abenchuchan Vita, ed.). 2014. The official ASL writing textbook. ASLized.org

Bébian R.A.A. 1825. Mimographie ou Éssai d'écriture mimique propre à régulariser le langage des sourds-muets. Éd. L. Colas (Paris). [www.archive.org/details/BebianMimographieLangageSourdsMuets1825]

Bianchini C.S., G. Gianfreda, A. Di Renzo, T. Lucioli, G. Petitta, B. Pennacchi, L. Lamano & P. Rossini. 2011. Écrire une langue sans forme écrite: réflexions sur l'écriture et la transcription de la Langue des Signes Italienne. *Travaux Linguistiques du CerLiCO* (Ed. PUR, Rennes), 24: 71-89.

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). Thèse, Université de Paris VIII et Università degli Studi di Perugia.

Bianchini C.S., F. Borgia & M. De Marsico. 2012. SWift: a SignWriting editor to bridge between deaf world and e-learning. *IEEE Proceedings of Intl Conference on Advanced Learning Technologies (ICALT)*: 526-530. DOI: 10.1109/ICALT.2012.235

Bianchini C.S. 2013. Rappresentare le Lingue dei Segni con SignWriting: evoluzione di SW sulla base delle esigenze di rappresentazione degli utenti. In : P. Fabbri, T. Migliore & A. Perri (eds) « Scritture per immagini », *Il Verri* (Milano), 53(oct): 126-135.

Bianchini C.S., F. Borgia & M. De Marsico. 2018a. A concrete example of inclusive design: deaf-oriented accessibility. In : K.L. Norman & J. Kirakowski (eds) « *The Wiley Handbook of Human Computer Interaction* », vol. 2 (ch. 33): 731-756. DOI: 10.1002/9781118976005.ch33

Bianchini C.S., L. Chèvrefils, C. Danet, P. Doan, M. Rébulard, A. Contesse & D. Boutet. 2018b. Coding movement in sign languages: the Typannot approach. *Proceedings 5th Intl Conference on Movement and Computing (ACM MoCo'18)*, sect. 1 (#9): 1-8. DOI 10.1145/3212721.3212808

Bianchini C.S. 2021. How to improve metalinguistic awareness by writing a language without writing: sign languages and SignWriting. In: Y. Haralambous (ed.) « *Grapholinguistics in the 21st century* », *Grapholinguistics and its Applications* (Fluxus Ed., Brest), 5(2): 1039-1065. DOI: 10.36824/2020-graf-bian

Bianchini C.S., F. Borgia & M. Castelli. 2021. L'appropriation et les modifications du SignWriting (SW) de la part de locuteurs de Langue des Signes Italienne (LIS). In : « *Transcrire, écrire, formaliser, 2* », *Travaux Linguistiques du CerLiCO* (Éd. PUR, Rennes), 31: 99-116.

Bickford J.A & M. McKay-Cody. 2018. Endangerment and revitalization of Sign Languages 1. In : L. Hinton, L. Huss & G. Roche (eds) « *The Routledge Handbook of Language Revitalization* », chap. 24. DOI: 10.4324/9781315561271

Bonnal-Vergès F. 2008. *Sémiogenèse de la Langue des Signes Française (LSF)*. Éd. Lambert-Lucas (Limoges).

Borgia F. 2010. SWift: Sign Writing improved fast transcriber. *Tesi di laurea magistrale*, Università di Roma 1 Sapienza.

Borgia F. 2015. *Informatisation d'une forme graphique des Langues des Signes: application au système d'écriture de SignWriting*. Thèse Université Toulouse 2.

Boutet D. & B. Garcia. 2006. Finalités et enjeux linguistiques d'une formalisation graphique de la Langue des Signes Française (LSF). In : R. Sabria (ed.) « Les langues des signes (LS) : recherches sociolinguistiques et linguistiques », *Glottopol*, 7: 32-52.

Boutet D., P. Doan, C.S. Bianchini, C. Danet, T. Goguely & M. Rébular. 2018. Systèmes graphématiques et écritures des langues signées. In : J.-M. Klinkenberg & S. Polis (eds) « Signatures: (essais en) sémiotique de l'écriture », *Signata*, 9: 391-426. DOI: 10.4000/signata.1684

Boutet D., C.S. Bianchini, P. Doan, L. Chèvrefils-Desbollies, C. Thomas, M. Rébular, A. Contesse, C. Danet, J.-F. Dauphin, M. Réguer. 2020. Réflexions sur la formalisation, en tant que système, d'une transcription des formes des Langues des Signes: l'approche Typannot. *SHS Web of Conferences (EDP Sciences)*, 78, #11001: 1-15. DOI: 10.1051/shsconf/20207811001

Boyes-Braem P. 2012. Evolving methods for written representations of signed languages of the Deaf. In A. Ender, A. Leemann & B. Wälchli « *Methods in Contemporary Linguistics* ». Mouton De Gruyter: 411-438.

Chevrefils L., C. Danet, P. Doan, C. Thomas, M. Rébular, A. Contesse, J.-F. Dauphin & C.S. Bianchini. 2021. The body between meaning and form: kinesiological analysis and typographical representation of movement in sign languages. *Languages and Modalities*, 1(1).

Chomsky N. 1957. *Syntactic structures*. Mouton (The Hague/Paris).

CNRS (Centre National de la Recherche Scientifique). 2020. Plan données de la recherche du CNRS. [www.science-ouverte.cnrs.fr/wp-content/uploads/2021/01/Plaquette-Plan-Donnees-Recherche-CNRS_nov2020.pdf]

Coulmas F. 2013. Written and unwritten language. In: F. Coulmas (eds) « *Writing and Society: an introduction* ». Key topics in sociolinguistics: 39-59. Cambridge University Press. DOI:10.1017/CBO9781139061063.004 Cuxac C. 1983. *Le langage des sourds*. Éd. Payot (Paris).

Cuxac C. 2000. *La Langue des Signes Française (LSF) : les voies de l'iconicité*. Éd. Ophrys (Paris).

Cuxac C. & E. Antinoro Pizzuto. 2010. Émergence, norme et variation dans les langues des signes : vers une redéfinition notionnelle. In: B. Garcia & M. Derycke (eds) « *Sourds et langues des signes: norme et variations* ». *Langage et Société*, 131: 37-53.

Cuxac C. 2013. Langues des signes : une modélisation sémiologique. *Nouvelle Revue de l'Adaptation et de la Scolarisation (INSHEA)*, 64: 65-80.

Cuxac C. & D. Boutet. 2014. *Corpus LS-Colin sur plusieurs genres discursifs (par Laurent Vallo, Nasredine Chab et Anthony Guyon) : version 1*. Huma-Num. DOI: 10.34847/COCOON.6D4B96B9-83C2-38B4-8FD6-ACB62E57DE4D

Danet C., D. Boutet, P. Doan, C.S. Bianchini, A. Contesse, L. Chèvrefils, M. Rébular, C. Thomas & J.-F. Dauphin. 2021. Transcribing sign languages with Typannot: a typographic system which retains and displays layers of information. In : Y. Haralambous (ed.) « *Grapholinguistics in the 21st century* », *Grapholinguistics and its Applications (Fluxus Ed., Brest)*, 5(2): 1009-1037. DOI: 10.36824/2020-graf-dane

Doan P., D. Boutet, A. Contesse, C.S. Bianchini, C. Danet, M. Rébular, J.-F. Dauphin, L. Chèvrefils, C. Thomas & M. Réguer. 2019. Handling sign language handshapes annotation with the Typannot typefont. In : « *Corpora and Representativeness* », *CogniTextes*, 19: 1-24. DOI: 10.4000/cognitextes.1401

DREES (Direction de la recherche, des études, de l'évaluation et des statistiques – République Française). 2007. Handicap auditif en France, apports de l'enquête HID 1998-1999. *Série Etudes*, 71. [www.drees.sante.gouv.fr/IMG/pdf/serieetud71.pdf]

- ELAN Team. 2002. ELAN:--Electronic Linguistic ANnotator. Max Planck Institute of Psycholinguistics, Nijmegen. [www.lat-mpi.eu/tools/elan/]
- Encrevé F. 2012. Les sourds dans la société française au XIX^e siècle : idée de progrès et langue des signes. Éd. Créaphis, Grâne.
- Frishberg, N. 1975. Arbitrariness and iconicity: historical change in American Sign Language. *Language*, 51(3): 696-719. DOI: 10.2307/412894
- Garcia B. 2004. Langue des Signes Française (LSF) : quelles conditions pour quelles formes graphiques ? In : A.M. Berthonneau & G. Dal (eds) « Linguistique de la LSF : recherches actuelles », *Sillexicales (Univ. Lille 3)*, 4: 173-183.
- Garcia B., J.-L. Brugeille, M.P. Kellerhals, A. Braffort, D. Boutet, P. Dalle & H. Mercier. 2007. LS-Script (2005-2007) : rapport final. Agence Nationale de la Recherche (ANR).
- Garcia B. & Sallandre M.A. 2013. Transcription systems for sign languages: a sketch of the different graphical representations of sign language and their characteristics. In : C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill & S. Tessendorf (eds) « Handbook of Body-Language-Communication », Mouton De Gruyter: 1125-1338. Gianfreda G., G. Petitta, C.S. Bianchini, A. Di Renzo, P. Rossini, T. Lucioli, B. Pennacchi & L. Lamano. 2009. Dalla modalità faccia-a-faccia ad una lingua scritta emergente: nuove prospettive su trascrizione e scrittura della Lingua dei Segni italiana (LIS). In : C. Consani, C. Furiassi, F. Guazzelli & C. Perta (eds) « Atti IX Congresso Internazionale dell'Associazione Italiana di Linguistica Applicata [AItLA] ». Guerra Ed. (Perugia): 413-437.
- Hopkins J. 2008. Choosing how to write Sign Language: a sociolinguistic perspective. *International Journal of the Sociology of Language*, 192: 75-89. DOI: 10.1515/IJSL.2008.036
- Hovy E. & J. Lavid. 2010. Towards a science of corpus annotation: a new methodological challenge for Corpus Linguistics. In : M.J. Blasco-Mayor & J.M. Martínez (eds.), *International Journal of Translation*, 22(1): 13-36.
- Johnston T.A. 2010. From archive to corpus: transcription and annotation in the creation of signed language corpora. *International Journal of Corpus Linguistics*, 15(1): 106-131. DOI: 10.1075/ijcl.15.1.05joh
- Klima E. & U. Bellugi (eds). 2008. *The signs of language*. Harvard University Press.
- Meletis D. 2020. *The nature of writing: a theory of grapholinguistics. Grapholinguistics and its Applications* (Fluxus Ed., Brest), 3: 473 p. DOI: 10.36824/2020-meletis
- MESRI-CoSO (Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation – Comité pour la Science Ouverte). 2018. *Plan national pour la Science Ouverte*. [www.ouvrirlascience.fr/plan-national-pour-la-science-ouverte]
- Millet A. 2006. *Le jeu syntaxique des proformes et des espaces dans la cohésion narrative en LSF*. Glottopol, 7. Millet A. 2019. *Grammaire descriptive de la langue des signes française : dynamiques iconiques et linguistique générale*. UGA Éd. (Grenoble).
- Minguy A. 2009. *Le Réveil Sourd en France : pour une perspective bilingue*. Éd. L'Harmattan (Paris).
- Mis J.-M. 2019. Reconnaissance de la langue des signes comme langue officielle de l'État français. Question # 21067 au Gouvernement Français, *Journal Officiel* 02/07/19: 5977 (réponse J.O. 14/01/2020: 258). <https://questions.assemblee-nationale.fr/q15/15-21067QE.htm>

- Mottez B. 2006. À propos d'une langue stigmatisée : la Langue des Signes. In : A. Benvenuto (eds.) « Les sourds existent-ils ? ». L'Harmattan, Paris.
- Newkirk D. 1989. SignFont handbook. Edmark Corp. (Bellevue WA).
- Perini M. 2013 Que peuvent nous apprendre les productions écrites des sourds ? Analyse de lectures écrites de personnes sourdes pour une contribution à la didactique du français écrit en formation d'adultes. Thèse, Université de Paris VIII.
- Pizzuto E. & P. Pietrandrea. 2001. The notation of signed texts: open questions and indications for further research. In : B. Bergman , P. Boyes-Braem, T. Hanke & E. Antinoro Pizzuto (eds.) « Sign transcription and database storage of Sign information », *Sign Language Linguistics*, 1/2 (spec. vol.): 29-43.
- Pizzuto E., P. Rossini & T. Russo. 2006. Representing signed languages in written form: questions that need to be posed. In : C. Vettori (ed.) *Proceedings of the 2nd Workshop on Representation and Processing of Sign Languages, 5th International Conference on Language Resources and Evaluation (LREC2006)*, Genoa 25/05/2006: 1-6.
- Prillwitz S., R. Leven, H. Zienert, T. Hanke & J. Henning. 1989. *Hamburg Notation System for Sign Languages: an introductory guide, HamNoSys version 2.0*. Signum Press, Hamburg.
- Repubblica Italiana. 2021. Conversione in legge, con modificazioni, del decreto-legge 22 marzo 2021, n. 41, recante misure urgenti in materia di sostegno alle imprese e agli operatori economici, di lavoro, salute e servizi territoriali, connesse all'emergenza da COVID-19. *Gazzetta Ufficiale* n. 120 (Suppl. 21L), Legge 21 maggio 2021 n. 69.
- Sallandre M.A. 2003. À la recherche des unités de la langue. Thèse, Université de Paris VIII.
- Sallandre M.A. 2014. Compositionnalité des unités sémantiques en langues des signes: perspective typologique et développementale. Habilitation à diriger des recherches, Université Paris VIII.
- Supalla S & L. Blackburn. 2003. Learning how to read and bypassing sound. *Odyssey*, 5(1): 50-55.
- Stokoe W.C. 1960. Sign language structure: an outline of the visual communication systems of the American deaf. *Studies in Linguistics*, 8 (occasional papers). [reprinted in: 2005, *Journal of Deaf Studies and Deaf Education*, 10(1): 3-37].
- Sutton V. 1995. *Lessons in SignWriting: textbook & workbook*. Deaf Action Committee for SignWriting, (La Jolla CA).
- Sutton V. 2010. SignMaker. [www.signbank.org/signpuddle/]
- UNESCO. sans date. Questions fréquentes sur les langues en péril. [www.unesco.org/new/fr/culture/themes/endangered-languages/faq-on-endangered-languages/]
- Union Européenne. 2016. Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016, relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données).
- Vermeerbergen M . 2006. Past and current trends in sign language research. *Language & Communication*, 26(2): 168-192.
- Volterra V. (ed.). 1987 [édition 2004]. *La lingua italiana dei segni*. Il Mulino (Bologna).
- Wilkinson M., M. Dumontier, I. Aalbergsberg, *et al.* 2016. The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, 3: #160018. DOI: 10.1038/sdata.2016.18

Woll B. 2013. The history of sign language linguistics. The Oxford handbook of Linguistics. Oxford University Press.

NOTES

1. L'analyse de SW a été réalisée dans le cadre de mon doctorat en cotutelle entre l'Université de Paris 8 et l'Università degli Studi di Perugia, et a été menée auprès de l'ISTC-CNR de Rome sous la supervision de Elena Antinoro Pizzuto. Un remerciement particulier va aux membres du groupe de recherche qui ont participé, il y a 10 ans, à la production et à l'analyse des textes en SW (A. Di Renzo, T. Lucioi, G. Gianfreda, L. Lamano, B. Pennacchi, P. Rossini, G. Petitta). La recherche et le développement du système de transcription typographique Typannot, encore en cours, est mené par l'équipe GestualScript de l'École Supérieure d'Art et Design d'Amiens (ESAD-Amiens), composée de chercheurs de l'ESAD, du laboratoire FoReLLIS de l'Université de Poitiers et du laboratoire DyLIS de l'Université de Rouen. Le projet est financé par l'ESAD, la Région Hauts-de-France, la DGCA du Ministère de la Culture, et a reçu par le passé des financements de la DGLFLF et du Mécénat du Crédit Agricole. Je remercie I. Marbot et M.L. Bianchini pour leur relecture attentive et M. Rébulard de l'aide fournie pour la réalisation des Fig. 10 et 11, ainsi que les relecteurs anonymes dont les remarques m'ont permis d'améliorer mon article et ma réflexion.

2. Ces règles sont liées non seulement à l'orthographe mais aussi à la disposition des caractères dans l'espace graphique, à la relation entre caractères, etc.

3. Dans cette optique, connaître les règles qui régissent le fonctionnement du Hangul ne font pas de l'auteur de cet article un scripteur du Coréen, tout comme le fait de maîtriser l'alphabet latin utilisé pour écrire le Finnois n'en font pas un scripteur du Finnois. En revanche, la connaissance de l'alphabet latin, associé à la connaissance du Français et de l'Italien, et de leurs conventions d'écriture respectives, en font un scripteur de ces deux langues.

4. Il importe de noter qu'il n'y a pas une seule et unique LS internationale et qu'il existe même des variantes régionales au sein d'un même pays (par exemple, il sera question au §2 de la variante romaine de la LS italienne). Dans cet article, l'auteur prend donc le parti de parler de LS au pluriel lorsqu'une affirmation peut être reconduite à l'ensemble des LS du monde, et d'utiliser les sigles identifiants les différents LS (LIS - LS italienne ; ASL - LS américaine ; LSF - LS française ; etc.) lorsqu'une affirmation se réfère à une ou plusieurs LS spécifiques.

5. En France, on estime que la LSF est pratiquée par environ 44 000 personnes sourdes, auxquelles il faut rajouter 75 000 entendants qui l'utilisent pour des raisons familiales et/ou professionnelles (DREES, 2007 ; p.78).

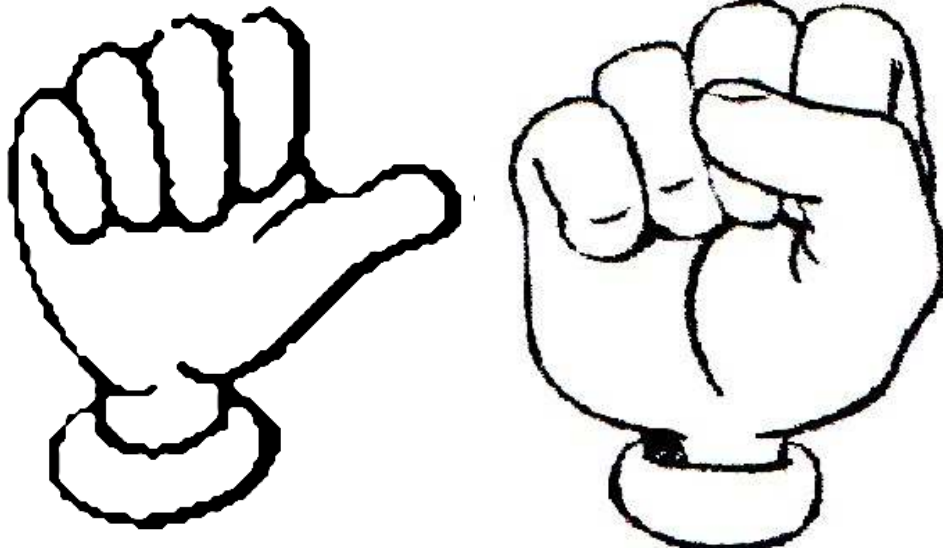
6. Serait-il possible d'adopter un système de représentation logographique ? Il faudrait pour cela associer un dessin à chaque référent que l'on voudrait écrire en LS. L'économie linguistique, qui s'applique aussi à l'écriture, ne permet pas la survie de systèmes purement logographiques : pour cette raison, même les systèmes d'écriture généralement considérés comme logographiques (comme les caractères chinois) ont tous une forte composante phonographique.

7. Dans les études sur les LS et la surdité, l'utilisation du mot « Sourds », avec un S majuscule, acquiert une valeur sociologique et culturelle. Cf. Mottez (2006).

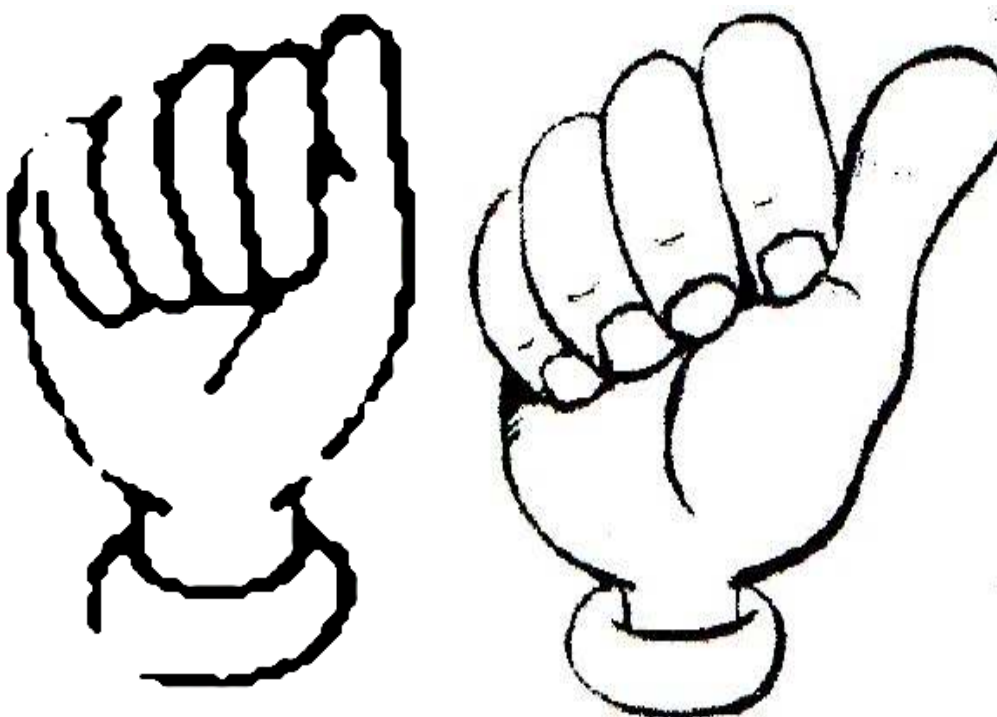
8. La reconnaissance des LS, et donc les politiques visant leur sauvegarde, varient encore énormément d'un pays à l'autre. En France, la Loi 2005/102 (Assemblée Nationale - Sénat, 2005) reconnaît la LSF comme « langue à part entière » mais l'insertion de la LS dans la Constitution n'est pas à l'ordre du jour (Mis, 2019), alors qu'elle est déjà une réalité en Autriche, Finlande, Hongrie, Afrique du Sud, Équateur, Kenya, Nouvelle-Zélande et Ouganda. En Italie, le premier acte législatif à caractère national donnant une reconnaissance à la LIS comme langue à part entière date du 21 mai 2021 (Repubblica Italiana 2021 ; art. 34-ter).

9. Aux alentours de 2008, Elena a décidé de modifier sa signature académique, passant de Elena Pizzuto à Elena Antinoro Pizzuto : selon la date des articles, il est donc possible de la trouver sous un nom ou sous l'autre. L'auteur de cet article fait le choix de se référer toujours à elle comme EAP, sauf dans les notices bibliographiques.
10. Le titre choisi pour cet article est un clin d'œil à celui-ci de EAP.
11. Une présentation, non exhaustive mais assez représentative, des différents systèmes existants pour représenter les LS est disponible sur le site <https://aslfont.github.io/Symbol-Font-For-ASL/>
12. Il est important de nuancer ce point : Brigitte Garcia, membre de l'équipe de Paris 8, s'intéresse au problème de la représentation graphique des LS mais d'un point de vue assez théorique. Elle recense en effet les caractéristiques que devrait avoir un système graphique spécifique aux LS pour remplir toutes les fonctions requises par un système d'écriture, sans toutefois proposer de solutions pratiques pour écrire les LS. De ses travaux (Garcia, 2004 ; Boutet et Garcia, 2006) vient l'idée que résoudre le problème de la représentation graphique de la LSF permettrait de résoudre la question pour toutes les LS.
13. « Étiquettes verbales » est le terme choisi par EAP pour décrire les étiquettes d'annotation en LV qui, dans la littérature sur les LS, sont normalement appelées « gloses ». La distinction qui doit être faite entre les « vraies gloses » (étiquettes d'annotation légitimement utilisées pour garder trace de l'analyse et indissolublement liées aux données) et les « gloses » telles qu'elles sont utilisées par les spécialistes de LS, est traitée au § 2.
14. D'où la possibilité, selon Millet (2019), d'attribuer une glose nominale ou verbale au même signe, choix opéré sur la base de la valeur « sentie comme plus centrale » (p.78) par le chercheur.
15. Le lecteur peut consulter le dictionnaire en ligne Spread the Sign pour satisfaire son éventuelle curiosité sur la forme et le mouvement de ce signe : <https://media.spreadthesign.com/video/mp4/10/173716.mp4>
16. La transcription phonétique des LV n'est pas une représentation exhaustive d'une production langagière : elle ne permet pas, par exemple, de noter les traits prosodiques, le rythme, etc. Elle permet toutefois de noter de façon assez précise l'ensemble des phonèmes produits.
17. Pour prendre connaissance de ces formes, le lecteur curieux peut consulter la section « note finale ».

18. Afin de satisfaire l'éventuelle curiosité du lecteur voici le S en LIS et en LSF



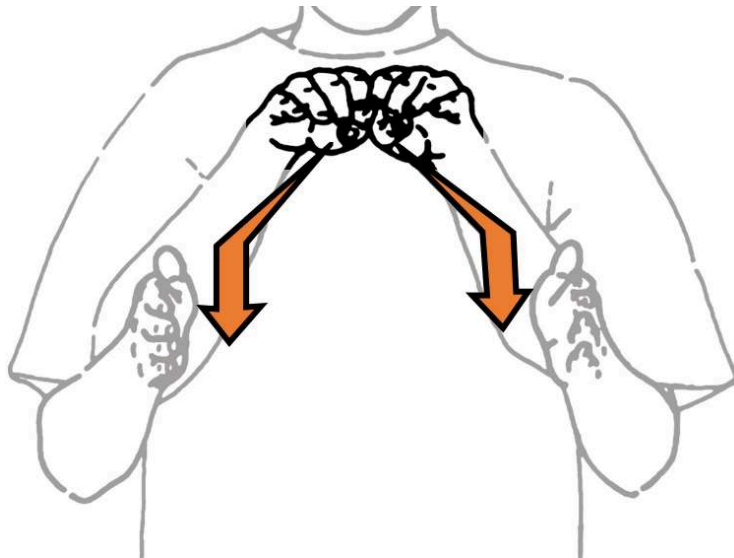
et le A en LIS et en LSF



19. Accessibles à l'adresse : <https://hal.archives-ouvertes.fr>

20. Pour prendre connaissance de ces formes, le lecteur curieux peut consulter la section « note finale ».

21. Afin de satisfaire l'éventuelle curiosité du lecteur voici le dessin du signe [MAISON] en ASL



Les parties du corps grisées ou absentes ne sont pas représentées dans la transcription fournie dans la Fig.3.

22. Le site <http://www.signwriting.org/> recueille la quasi-totalité des manuels, du matériel et des articles produits sur et en SignWriting depuis son invention.

23. Même sans connaître SW, dans la Fig. 5 il est possible de deviner aisément les expressions faciales (en vert), les formes de la main (en bleu), les bras et les épaules (en noir) et les directions et ampleurs des mouvements des mains (flèches rouges) et de la tête (flèches vertes). En revanche, la compréhension des contacts (astérisques et croix rouges), des mouvements des doigts (points rouges) et les indications dynamiques (tildes et parenthèse rouges) est moins intuitive.

24. Cette manière de faire est aussi utilisée dans les IDgloss (Johnston, 2010), qui sont des étiquettes verbales dont le but est d'identifier (sans décrire) la forme des signes, sans nécessairement faire référence à leur signification. Pour faire cela, une IDgloss (souvent une traduction en LV du signe, mais qui pourrait très bien être un numéro progressif dépourvu de sens) est associée à tous les signes ayant (ou semblant avoir) une même forme. Si la forme des signes a été transcrite (souvent avec HamNoSys), il sera possible, en requêtant une IDgloss, d'obtenir une liste de signes que le transcripateur a jugé identiques malgré des nuances dans la manière dont ils ont été transcrits.

25. Fabrizio Borgia avait proposé de nouveaux logiciels permettant une meilleure expérience utilisateur et surtout une accélération du processus d'écriture de SW mais ces logiciels — fruits de ses travaux de master (Borgia, 2010) et de thèse (Borgia, 2015) — n'ont pas été diffusés auprès du grand public. Pour plus de détail cf. Bianchini *et al.* (2012 ; 2018a).

26. Il faut souligner, en toute honnêteté, que les logiciels de traitement de texte ne prévoient pas la possibilité d'une écriture non linéaire comme SW : l'absence d'une police fonctionnelle est donc plus probablement due à des contraintes techniques qu'à un manque de volonté de la part de l'équipe de Sutton.

27. Cette liste concerne les systèmes de transcription des LS ; pour un système d'écriture, les critères seraient différents mais les déterminer n'est pas l'objet de cet article.

28. Ce concept est assez subjectif ; toutefois, l'auteur considère qu'avoir mis presque une heure pour coder les six exemples présents dans la Fig. 5, ne fait pas de SW un système très « rentable », surtout au vu de ses limites en matière de requêtabilité.

29. Dans les articles de l'équipe GestualScript, les concepts de requêtabilité, rentabilité et expansibilité sont recouverts par ceux de « modularité » et de « généricité » (Bianchini *et al.*, 2018b).

30. Unicode est le standard assurant l'interopérabilité informatique des systèmes d'écriture, quel que soit le système d'exploitation, le logiciel ou le navigateur utilisé.

31. Les glyphes choisis ici partagent les propriétés graphiques permettant au lecteur d'identifier un A majuscule (deux lignes se rejoignant au sommet, traversées en leur milieu par une ligne horizontale). Il serait toutefois possible d'associer aussi à U+0042 un glyphe ne ressemblant pas à un A – comme [%], [Z] ou même le dessin d'un dinosaure – qui rendrait le texte illisible pour l'humain mais non pour les logiciels, qui lisent les caractères et non leurs manifestations glyphiques.

32. L'absence d'un glyphe correspondant à ce caractère dans le Tab. 3 n'est pas un oubli : ce caractère est labélisé ZERO WIDTH JOINER et, en tant qu'espace sans largeur, il ne peut apparaître à l'écran ; il existe donc des caractères qui ne peuvent être associés à un glyphe mais qui influent les glyphes de caractères se ligaturant avec eux.

33. L'absence d'un glyphe correspondant à ce caractère dans le Tab. 3 n'est pas un oubli : ce caractère est labélisé ZERO WIDTH JOINER et, en tant qu'espace sans largeur, il ne peut apparaître à l'écran ; il existe donc des caractères qui ne peuvent être associés à un glyphe mais qui influent les glyphes de caractères se ligaturant avec eux.

34. Le processus d'acceptation d'un système graphique de la part du Consortium Unicode est long et complexe et ne peut être lancé qu'une fois que le système est considéré complet et stable, ce qui n'est pas encore le cas de Typannot.

35. L'association des données de MoCap avec Typannot et leur implémentation dans le clavier virtuel est encore au stade embryonnaire et constitue, entre autres, l'un des résultats attendus des thèses de Léa Chevretil et Chloé Thomas, doctorantes de l'Université de Rouen et membre de l'équipe GestualScript.

36. Pour ceux qui ne connaissent pas le travail des dessinateurs de caractères, le fait de dessiner les glyphes composés peut sembler une opération anodine, se résumant à tracer des silhouettes de mains sur du papier. Il s'agit toutefois d'une opération beaucoup plus complexe, qui nécessite un réel travail de recherche en design graphique, le but étant de construire des modules qui puissent s'agencer de façon automatique et assurer la lisibilité de chaque configuration existante ou réalisable (le chiffre s'élève à plusieurs millions). Pour plus de détail sur cette phase, cf. Boutet *et al.*, 2020 ; Danet *et al.*, 2021.

37. En dehors des travaux de EAP et de ses collaborateurs, il ne nous semble pas que la critique des gloses ait été traitée par des chercheurs comme sujet central de leur recherche. Toutefois, des auteurs comme Vermeerbergen (2006), Boyes-Braem (2012), Woll (2013) ou encore Garcia et Sallandre (2013) ont souligné les problèmes liés à l'utilisation des gloses, se référant souvent aux travaux de EAP.

ABSTRACTS

Sign Languages (SLs) are languages which do not have a writing system of their own and which, by their visual-gestural nature, cannot be written phonographically. The consequences of this "unscriptability" are at the center of the work of Elena Antinoro Pizzuto (EAP), carried out from

2000. EAP, who died in 2011, highlighted the biases that glosses, used as substitutes for transcription, induce in the linguistic analysis of SLs. Over the past 10 years, technological and methodological advances have made it easier to automatically enter data from the body (thanks to motion capture), to annotate multilinearly multimodal corpora and to maintain the links between annotations and videos (thanks to software such as ELAN), as well as to distribute videos on which the analyses are based (thanks to F.A.I.R. principles). However, these technologies and good practices — although fundamental — do not act as a transcription system: they are solutions that eliminate some of the symptoms linked to SLs unscriptability, but do not make up for the lack of a system of graphic representation for these languages. Therefore, the questions raised by EAP are still relevant today. While various attempts to transcribe SLs have emerged, their use to exploit large corpora of SLs has often been hindered by their poor readability and/or queryability. By comparing the pros and cons of these systems, it is possible to draw up a list of characteristics that a transcription system for SLs (and for co verbal gestures too) should have. Nevertheless, developing a system that brings together all these specificities is a challenge that linguists cannot meet alone: it is in multidisciplinary, and particularly in working jointly with specialists in the graphic aspects of writing, that the solution lies. In fact, the GestualScript team has employed this "grapholinguistic" approach to develop Typannot, a typographic transcription system for SLs and co-verbal gestures.

Les langues des signes (LS) sont des langues qui n'ont pas un système d'écriture qui leur est propre et qui, par leur nature visuo-gestuelle, ne peuvent pas être écrites phonographiquement. Les conséquences de cette « non-scriptibilité » sont au centre des travaux menés à partir de 2000 par Elena Antinoro Pizzuto (EAP). EAP, décédée en 2011, mettait en évidence les biais que l'utilisation des « gloses », employées comme substituts de la transcription, induisent dans l'analyse linguistique des LS. Ces 10 dernières années, les avancées technologiques et méthodologiques ont rendu plus aisée la saisie automatique des données corporelles (grâce à la capture de mouvement), l'annotation multilinéaire de corpus multimodaux et le maintien des liens entre les annotations et les vidéos (grâce aux logiciels comme ELAN), ainsi que la diffusion des vidéos sur lesquelles se basent les analyses (grâce à l'application des principes F.A.I.R.). Toutefois, ces technologies et bonnes pratiques — bien que fondamentales — ne font pas office de système de transcription : ce sont des solutions qui suppriment une partie des symptômes liés à la non-scriptibilité des LS, mais ne combrent pas l'absence d'un système de représentation graphique de ces langues. Les questions soulevées par EAP sont donc encore d'actualité. Si différentes tentatives pour transcrire les LS ont vu le jour, leur utilisation pour exploiter de vastes corpus de LS s'est souvent heurtée à leur faible lisibilité et/ou requêtabilité. En comparant les avantages et les inconvénients de ces systèmes, il est possible d'établir une liste des caractéristiques que devrait posséder un système de transcription des LS (et de la gestualité co-verbale). Mais développer un système qui réunit toutes ces spécificités est un défi que les linguistes ne peuvent pas relever seuls : c'est dans la pluridisciplinarité, et en particulier dans le travail conjoint avec les spécialistes des aspects graphiques de l'écriture, que semble résider la solution. Cette approche, dite « grapholinguistique », est suivie par l'équipe GestualScript pour développer Typannot, un système de transcription typographique des LS et de la gestualité co-verbale.

INDEX

Keywords: sign language, transcription, annotation, multimodal corpora, grapholinguistics

Mots-clés: langue des signe, transcription, annotation, corpus multimodaux, grapholinguistique

AUTHOR

CLAUDIA S. BIANCHINI

Université de Poitiers, Unité de Recherche 3816 FoReLLIS



Corela

Cognition, représentation, langage

2021

Vol. 19, n°2

Représenter les langues des signes sous forme écrite : questions qui ont besoin (encore aujourd'hui) d'être posées

Claudia S. Bianchini



Electronic version

URL: <https://journals.openedition.org/corela/13190>

DOI: 10.4000/corela.13190

ISSN: 1638-573X

Publisher

Cercle linguistique du Centre et de l'Ouest - CerLICO

Electronic reference

Claudia S. Bianchini, "Représenter les langues des signes sous forme écrite : questions qui ont besoin (encore aujourd'hui) d'être posées", *Corela* [Online], | 2021, Online since 02 December 2021, connection on 06 December 2021. URL: <http://journals.openedition.org/corela/13190> ; DOI: <https://doi.org/10.4000/corela.13190>

This text was automatically generated on 6 December 2021.



Corela – cognition, représentation, langage est mis à disposition selon les termes de la licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Partage dans les Mêmes Conditions 4.0 International.