e-LIS: Electronic Bilingual Dictionary Italian Sign Language-Italian

Chiara Vettori
Mauro Felice
European Academy of Bozen/Bolzano

This paper presents the design of e-LIS (Electronic Bilingual Dictionary Italian Sign Language (LIS) - Italian), an ongoing research project at the European Academy of Bolzano started in 2004. It is the first attempt to build a sign-language reference dictionary that contains definitions and examples in the sign language itself and which offers a search engine that guides the user in the process of reconstructing and retrieving the sign he is looking for. Thus, not only can users go from Italian to LIS, but also from LIS to Italian.

1. Introduction

Around 969,000 people in Italy are deaf.[1] The first language of the majority of them is LIS, Lingua Italiana dei Segni (Italian Sign Language). LIS is also acquired as a second or third language by hearing family members, teachers, interpreters and logopedics, amounting to far more than one million people using LIS, in various degrees of language competence. Unfortunately, the quality and accessibility of LIS-courses and supporting material (dictionaries, grammar books, text books etc.) lack behind the actual need. As for LIS dictionaries, the vast majority of them are paper based ones, e.g. Radutzky 1992 (752 signs, 2,500 sign meanings); Angelini et al. 1991 (400 signs), Romeo 2004 (1,300 "images" and 150 sentences in LIS). The paper format cannot obviously account for the possibility of describing the three-dimensional, mobile complexity of each sign. A first, significant attempt in Italy to exploit new technologies to approach sign languages in an innovative and more proficient way, was made by the team of Cooperativa Alba in Turin. Its members have created an Internet portal for LIS (DIZLIS) that now features more than 1,000 video-filmed signs, which represent a respectable size for a sign language dictionary, cfr. Sternberg 1987 (3,300 signs).

2. Towards e-LIS

Most sign language dictionaries form a hybrid between a reference dictionary and a learner's dictionary. This often occurs because sign language is implicitly considered as the second language of a "learner's dictionary" de-facto created for the needs of hearing people. At the same time these lexicographic works pretend to fulfill the function of a reference dictionary of the involved sign language, only in virtue of the presence of drawings and/or photos representing different signs. "A major feature of such dictionaries is the absence of definitions, it being assumed that each sign would have exactly the same meaning(s) as the written word with which it is linked" (Brien 1997). This sort of production treats signs as equivalents of the words of a spoken language and neglects the complexity, the dignity of sign language and its peculiarities in semantics and syntax.

As we assume that LIS is an autonomous language and not a mere visual representation of Italian, we designed a dictionary, e-LIS, which describes two systems at the same time, the Italian and the LIS one, and which can also build a bridge between them through a sort of "translating interface". In this perspective, accepting Stokoe's description of what he calls "serious (sign language) dictionaries" (Brien 1997), we are greatly motivated to focus on the definition of sign meanings that could reveal much of the deaf culture.

[1] 92,000 of them are pre-lingual deaf (information update 2000).

Starting point for each lexicographic work is the building of a corpus of the target language. No corpus of LIS is actually available[2] so we had to start from previous lexicographic experiences in LIS to establish the list of "basic" signs to include in e-LIS. Obviously enough, as there is no corpus of LIS, there are no frequency studies that give a range to each sign and from which to determine the level of elementariness/complexity of a sign. Although we are aware that it was a straining, we decided to assign a range to the signs we collected from the DIZLIS project and from the dictionary of LIS already published in our country by comparing them with the many frequency lists available for spoken Italian (De Mauro et al. 1993, etc.). It resulted a list of signs which do appear to pertain to an elementary vocabulary, but we need further confirmation from the informants who help us in deciding about problematic cases [such as i.e. the word/sign "name" which has a low range (= high frequency) in Italian but which could have a much higher level in sign language where the concept of "name" is almost substituted by a sign that identifies the person it refers to].

3. e-LIS: structure and search engines

The e-LIS dictionary contains two sections respectively labelled ITA-LIS and LIS-ITA. Both sections offer a variety of lexicographic information about the word/sign the user is looking for. In the LIS-ITA module the sign, reproduced through a video, is explained and exemplified by a definition and a sentence which both occur in LIS. The translation in Italian, imprescriptable as it is a bilingual dictionary, is a paraphrase apt to (better) explain the concept expressed in LIS for those users who need it in Italian. Viceversa, in the ITA-LIS module the lemma is defined and explained in Italian and accompanied by a translation/explanation in LIS. This is obviously meant to give voice to both cultures: what counts in the Italian deaf culture may do not count in the Italian one. That's why we decided for the two parallel sections which do communicate one with the other, but which could also stand alone. This independence underlines the willingness of our team to acknowledge LIS as an autonomous language—in respect to Italian—with own metalinguistic properties.

[2] The Institute of Cognitive Sciences and Technologies (ISTC) of the National Research Council (CNR) in Rome has collected 200 hours of sign language recordings that have been neither trascripted nor properly archived yet. (http://*www.let.kun.nl/sign-lang/echo/docs/SL_corpora_overview.pdf*, SL_corpora_overview, ECHO project 2003).

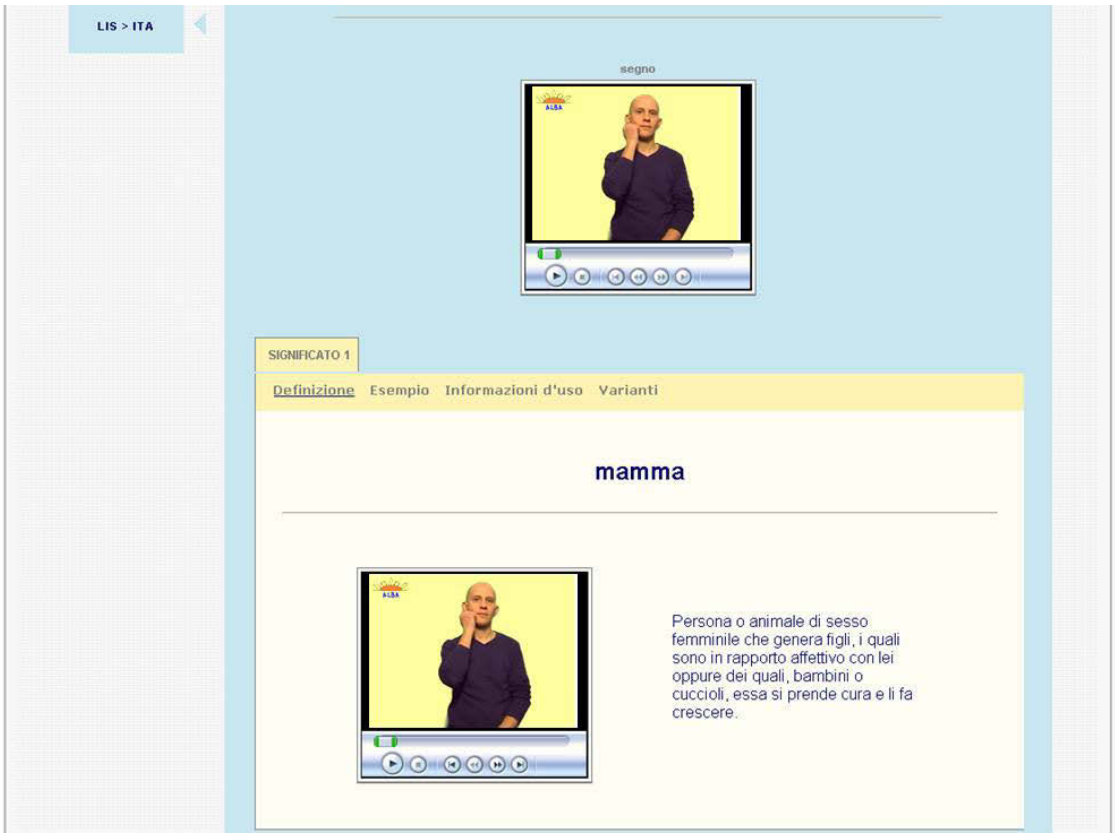

Figure 1. e-LIS, LIS-ITA section.

It is now clear that e-LIS accommodates for two distinct user groups. (a) Hearing Italian people who study LIS and who will start with an Italian query term in an Italian environment (Italian definitions, explanations etc.); (b) LIS-signers looking for a sign and who should have the possibility to formulate query terms in LIS and have a LIS environment. To realize this latter possibility, we have implemented two different search engines. The ITA-LIS one does not contain innovative features with respect to other online dictionaries for any verbal languages: users can type the word they are looking for and the system provides a result set. Besides the "classical" search service, users can look for one word using an alphabetical index, which groups the words depending on their first letter. As for the LIS-ITA search engine, it represents the most innovative concept of the whole project, both from a linguistic and from a technical point of view as it is the first attempt to build a systematic and intuitive procedure to look up a specific sign in the dictionary.

Considering that a sign is a combination of four so-called parameters (cfr. the studies of Stokoe and others), *handshape*, *palm orientation*, *location* and *movement*[3], we have conceived a search engine that will accompany the user in the process of deconstructing and reconstructing a sign. In order to represent graphically the structure of a sign, we analyzed different transcription systems[4] and decided to use the Stokoe-based one adapted for LIS by Radutzky, who employed it in her dictionary and which is based on a series of symbols and icons that combine in a string "describing" a sign which is not immediately decipherable, nor transparent. We transformed these values that constitute the four parameters (i.e. there are 57 values for handshape, 6 for

[3] There's also a fifth parameter, that of non-manual features, "a series of meaningful units in sign language which convey additional information concomitant with the meaning being expressed by the hand shapes" (O'Baoill and Matthews 2000: 167).

[4] Transcription systems are "attempts to represent signed languages on the level of formational components such as those first isolated by Stokoe. These systems make use of sets of iconic symbols for handshapes, locations, movements, and nonmanual elements, and provide special keyboards and related computer facilities" (Hoiting and Slobin 2002). The best known are the Stokoe-based one, SignWriting and HamNoSys.

orientation, 16 for location and ca. 30 for movement) in ASCII format so that the machine could read them and represent them through a series of drawings and animated gif which should result transparent to the user who's looking up a sign. The e-LIS dictionary offers a wizard-like interaction: the user can select one parameter at a time (e.g. *handshape*) and choose its value among all the values of the chosen parameter (e.g., all specific handshapes) that the system proposes to him/her (see Figure). Once the user has chosen a value, this element becomes a search parameter. A visual feedback is shown on the left side of the screen, i.e. the user sees a miniature of the value he/she has just selected. After that, the user can choose to either trigger the search engine or to set another parameter value up to the maximum of all four parameters. It is worth noting that during the reconstruction/definition of a sign, users interact only with iconic intuitive images representing sign parameters, while the encoding of the transcripted signs remains hidden, thus increasing the user-friendliness of the system, allowing even non experts of sign transcription systems to profitably use the dictionary.

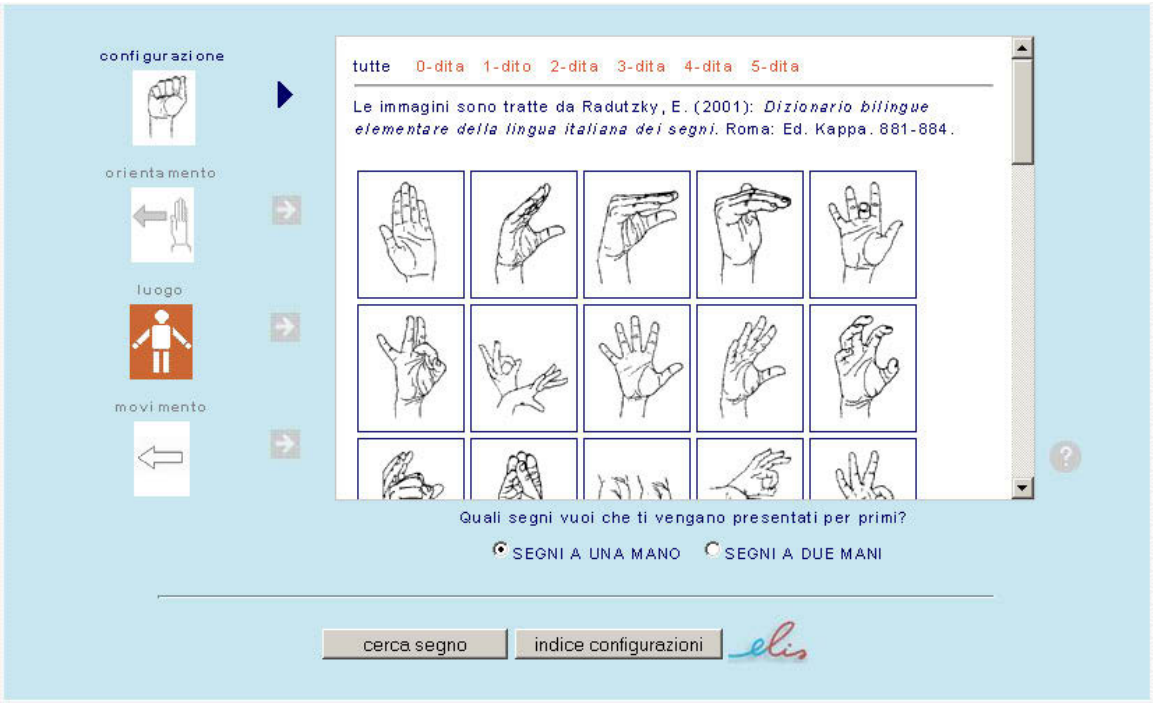

Figure 2. e-LIS, the interface for the specification of the sign (in this case, all four parameters have been specified by the user)

Once the system has retrieved all the signs stored in the e-LIS database matching the input sign, the first sixteen results (eight per page) are shown via a snapshot representing a real person in the act of signing. Users can select their preferred result and go to the lexicographic card file by clicking on it. Here, a video shows a real person in the act of reproducing that specific sign. Moreover, definition, examples, usage information and variants are shown by means of other videos. At this point, users can go back to the result set to visualize another sign, or they can start a new search session. The digital videos embedded in the dictionary pages improve the dictionary usability, and help making signs, examples and the other information immediately and simply comprehensible.

The sign search engine has an important feature: beyond the possible exact match, similar signs (with respect to the input one) are retrieved. This is achieved by using the Stokoe parameters as similarity criteria: given the input sign, the number and the type of equal parameters are used to rank all the signs in the e-LIS database (even if only sixteen signs are included in the results). In particular, given the input sign S, and given two signs S1 and S2, having respectively three and two parameters equals to S, we say that S1 is more similar to S than S2. As for the type, we decided to give parameters a relative order, i.e., from the most to the least relevant, *handshape*, *movement*, *location*, *palm orientation*. This order relies on a series of considerations and tests we made that permitted us to claim that, in our application domain, *location* is more ambiguous

than *movement* (because of the many signs which are signed in the so-called neutral space) so that the value referring to *location* parameter is excluded before the one relative to *movement*. The inclusion of similar signs in the result set allows to overcome two problems: on the one hand, the number of signs in the e-LIS database is not yet big, thus it is very presumably that users define a sign that has not been stored in the e-LIS database, so that the query returns no results. On the other hand, users cannot be assumed to exactly remember or know the sign they are searching for: they could make mistakes, they could choose a different handshape, they could try and define a sign with no specific purpose. Including similar signs allows offering a kind of did-you-mean Google™ behaviour, thus "forcing" the user to reflect on the sign structure.

In conclusion, we have illustrated e-LIS: a bilingual electronic dictionary Italian Sign Language (LIS) – Italian that represents the first attempt in Italy to build a reference LIS dictionary and to publish it on Internet so that a great audience can use it. Moreover, e-LIS testify our effort to draw attention to Italian deaf people, their culture and their natural language, LIS.

References

Angelini, N. et al. (1991). *I primi 400 segni. Piccolo dizionario della Lingua Italiana dei Segni per comunicare con i sordi*. Firenze: La Nuova Italia.

Brien, D.; Collins, J. (1997). "The CD-ROM Dictionary of Deaf Community and Culture: Some Issues in the Creation of Sign Language Dictionaries". Paper presented at *Empower'97: International Conference on Deaf Education* [online]. *http://www.ssc.mhie.ac.uk/archive/proceedings/briencollins.html*

De Mauro, T. et al. (1993). *Lessico di frequenza dell' italiano parlato*. Milano: EtasLibri.

Hoiting, N.; Slobin, D. I. (2002). "Transcription as a tool for understanding: The Berkeley Transcription System for sign language research (BTS)". In Gary, M.; Woll, B. (eds.). *Directions in Sign Language Acquisition*. Amsterdam: John Benjamins. 55-75.

O'Baoill, D.; Matthews, P. A. (2000). *The Irish Deaf Community*. Vol. 2: The Linguistics of ISL. Dublin: ITE.

Radutzky, E. (ed.) (1992). *Dizionario bilingue elementare della Lingua Italiana dei Segni*. Roma: Edizioni Kappa.

Romeo, O. (2004). *Il dizionario tematico dei segni*. Bologna: Zanichelli.

Sternberg, M. L. A. (ed.) (1987). *American Sign Language Dictionary*. New York: Harper & Row.