

A bilingual dictionary Mexican Sign Language-Spanish/Spanish-Mexican Sign Language

Antoinette Hawayek*, Riccardo Del Gratta⁺, Giuseppe Cappelli⁺

*Universidad Autónoma Metropolitana, ⁺Istituto di Linguistica Computazionale

*Iztapalapa / México, ⁺Consiglio Nazionale delle Ricerche

Pisa, via Moruzzi 1, Italia

*anha@xanum.uam.mx, ⁺riccardo.delgratta@ilc.cnr.it, ⁺beppe@ilc.cnr.it

Abstract

We present a three-part bilingual *specialized* dictionary Mexican Sign Language-Spanish / Spanish-Mexican Sign Language. This dictionary will be the outcome of a three-years agreement between the Italian “Consiglio Nazionale delle Ricerche” and the Mexican Conacyt. Although many other sign language dictionaries have been provided to deaf communities, there are no Mexican Sign Language dictionaries in Mexico, yet.

We want to stress on the *specialized* feature of the proposed dictionary: the bilingual dictionary will contain frequently used general Spanish forms along with scholastic course specific *specialized* words whose meanings warrant comprehension of school curricula.

We emphasize that this aspect of the bilingual dictionary can have a deep social impact, since we will furnish to deaf people the possibility to get competence in official language, which is necessary to ensure access to school curriculum and to become full-fledged citizens.

From a technical point of view, the dictionary consists of a relational database, where we have saved the sign parameters and a graphical user interface especially designed to allow deaf children to retrieve signs using the relevant parameters and, thus, the meaning of the sign in Spanish.

1. Introduction

Speech perception consists of a set of computations that take continuously varying acoustic waveforms as input and generate discrete representations that make contact with the lexical representations stored in longterm memory (Friederici, 2005; Alter and Friederici, 2004; Poeppel and Philip J., 2008).

In the process of language acquisition, language learners store the representations that will enter into subsequent linguistic computation. Deaf children, however, cannot perceive acoustic signals that are transformed into linguistic representations and, consequently, they can not acquire oral languages naturally.

Although it is universally accepted that deaf children have to acquire a signed language to intercommunicate, also a basic competence in official language is necessary to ensure access to school curricula and development of academic skills. They also need to understand written language to become full-fledged citizens.

Moreover, even in countries like USA and Sweden, where they have had special education for deaf children for over a century, the proficiency in the oral language is below expectations (Prinz, 1988; Ramsey, 1997).

About the situation in Mexico, we can affirm that al-

though the education for students with special needs is being restructured, no program for the teaching of Spanish as a second language has been developed, yet. Our evaluation of sixth graders written Spanish in three schools for deaf children in Mexico City shows that their linguistic competence is far from the minimum required to understand textbooks used in Mexican schools, as shown by research done at the Universidad Autónoma Metropolitana (Hawayek, 2006a; Maya Anguiano, 2002; Romero Alpizar, 2002; Taba Nambo, 2003).

The specialized bilingual dictionary Spanish-MSL¹ / MSL-Spanish we are going to present, will be the only tool available to Mexican deaf students to access the concepts contained in grammar school textbooks.

2. The dictionary

Sign languages “engineering” requires that sign parameters have been previously classified and analyzed before software engineers can work with stable information to retrieve and manage data in a faster manner. In order to define the bilingual dictionary, we have followed this as guideline to first catalog and classify sign parameters into a dedicated relational database and then to be able to define, besides this

¹The acronym MSL stands for Mexican Sign Language.

database, a graphical user interface (GUI) which can be used by deaf children to retrieve signs using the relevant parameters and, thus, the meaning of the searched signs in Spanish.

The proposed dictionary consists of two distinct sections, an MSL-Spanish and a Spanish-MSL, and a multimedia shared part which collects videos. Videos act as bridge between the two sections. These distinct sections are handled by means of a database, while videos are stored in a dedicated file system. On top of the database, we have defined a GUI, especially designed to access the data contained in the database. The GUI provides a search engine to end users (deaf children and teachers) for browsing both sections of the database. In addition, it provides, to sign language experts, an administrative interface designed to tag videos with forms and signs parameters (for more details on database cf. sections 2.1. 2.2. and 2.3.).

2.1. Spanish-Mexican Sign Language

The Spanish-MSL section of the dictionary is a standard form enumerated dictionary. It is semi-automatically created by a morphosyntactic tool for linguistic analysis, (cf. section 2.1.1.), and based upon Department of Education textbooks.

The goal of the Spanish-MSL section of this bilingual specialized dictionary is, thus, to allow Mexican deaf students to retrieve the MSL meaning of the words contained within the “Mexican Secretary of Education” textbooks used in all grammar schools in México. Being a form enumerated dictionary, the MSL Spanish section will allow users to find signs which decode the form they are looking for, without knowing the lemma to which it corresponds.

This feature, therefore, makes the dictionary accessible to users with almost no proficiency in Spanish and who would not be able to link, for example, the different forms of a verb to the only form (the infinitive) that appears in non enumerated Spanish dictionaries. Clearly, in designing this section of the dictionary, as there are no time or person marks in MSL, we took into consideration that various Spanish forms have to be mapped onto one single sign. In conclusion, in this section of the dictionary, many Spanish forms are linked to the MSL verb that adequately represents all the conjugated verb forms of a Spanish verb (cf. again section 2.1.1.).

2.1.1. Ayda

As a morphological marker, AyDA, (Saba and Cappelli, 2003), is based on a reference and an optional dictionary. The optional dictionary makes it possi-

ble for researchers to create a personal list of forms not contained in the reference dictionary. This is extremely important for extracting low frequency specialized words of areas such as natural science, mathematics, and Mexican languages words incorporated in Mexican Spanish.

The Department of Education school textbooks are semi-automatically analyzed thus the final list of extracted words will contain frequently used general forms along with specific lower frequency specialized words whose meanings warrant comprehension of school curricula. Table 1 displays an extract of the AyDA outcome.

Domain	Lemma	Form
General	tener	tengo tenía ... dado
	dar	dar ...
Specific	parar	para pares ... sobre
	sobrar	sobra ...

Table 1: Ayda outcome example

2.2. Mexican Sign Language-Spanish

The MSL-Spanish section is designed to hold the “alphabet” of MSL. To create the alphabet necessary to access signs, we have analyzed the four traditional parameter of signed languages (hand-shape, place of articulation, movement, hand orientation), (Brentari, 1998). As there are few analysis of MSL grammar, (Hawayek, 1997; Hawayek, 2006b; Hawayek, 2007; Hawayek and Cappelli, 2004; Hawayek and Capistrán, 2008) among others, we had to create the inventory of the entities belonging to each parameter. The inventory of the entities belonging to each parameter constitute by itself an “alphabet” that guides the retrieval of each sign.

2.2.1. Hand-shapes

To create the list of hand-shapes, we examined, besides MSL manual alphabet for finger spelling, other signed language alphabets, MSL classifiers provided other hand-shapes. For the users of this dictionary (grammar school children) the MSL hand-shapes have to function as the “alphabet” that constitutes the first

step in the retrieval of the signs that represent the concepts they want to map onto Spanish. Consequently, the list of MLS hand configurations have been classified and ordered in a way that can be naturally perceived by young children, the way pre school hearing children naturally recognize phonemes.

2.2.2. Place of Articulation

Signs are usually produced in the space in front of the signer at the level of the torso (*neutral space*) but many signs make contact with the body. As same hand-shape may acquire different meanings depending on place of contact, so it is necessary to identify all the points with which the signer may make contact. We have listed places of contact in head/face (such as, for instance, forehead, cheeks, ...) as well as in the body (such as neck, chest, stomach, arm, and wrist). As we show in section 2.2.6., we have simplified the searching engine by retrieving two-handed signs with the same mechanisms used to retrieve one-handed signs, so we have also listed the different parts of the non-dominant hand (palm, back of palm, and fingers) that function as a place of articulation.

2.2.3. Movement

Some phonological models propose that signed languages do not have movement features, that movements are the result of interpolation between points of stasis. This means that, for example, an arc movement is analyzed as having an orientation change and a location change.

An important goal is to select the minimum number of features required to identify each sign in the table containing all videos. It is more economical if movement feature is adopted; movement takes care of orientation change and location change that "movementless" models have to specify.

However, there are signs that select the same hand configuration and the same point of articulation but they are distinguished by a movement feature performed by the dominant hand while (or just after) making contact with the point of articulation, as reported below².

gray and lawyer These signs show the same hand configuration and select the same point of articulation but while in *gray*, after establishing contact, the dominant hand moves clockwise, in *lawyer* it sways to the right and to the left, see figure 1;

owner and God In these two signs, the dominant

hand selects the same hand-shape and makes contact with same place of articulation (temple) but while for *owner*, the hand moves away from the temple towards the front of the signer, in *God* it moves upwards, see figure 2.

If videos are tagged ranking the sign parameters according to the information they provide the result will be that only one sign (the one the user is trying to retrieve) will appear on the screen.

2.2.4. Orientation

Phonologists accept that the orientation parameter is quite complex to represent. We hypothesized that the movement parameter can frequently substitute the orientation parameter. Many of the MSL signs we have analyzed can be identified by the movement that causes a change in orientation. However, the orientation parameter provides essential information for the retrieval of signs like *affection* and *captain*. In fact, these two signs select the same hand configuration (five fingers, bent) and same point of articulation (opposite shoulder of dominant hand), but they differ by the orientation of the hand: when signing *captain*, palm of the hand faces the shoulder, tips of all fingers come in contact with shoulder; when signing *affection*, palm faces downward, so it is the edge of the hand what makes contact with the shoulder.

2.2.5. Two-handshape signs

In sign language, there may be a change of handshape in the articulation of a sign. Here by handshape change we mean the change of the handshape from one configuration to another³.

There are two types of handshape changes: handshape contours which are open or closed variations of the same handshape and handshape contrasts which specify contrastive hand-shapes by selecting different fingers, (Brentari, 1990; Brentari, 1998).

None of the signs with handshape contours that we have analyzed have presented any problems but, on the contrary there are high frequency words in MSL, originated by finger spelled forms, where the form has come to stand not for each letter of the Spanish word and most of these signs have two handshapes, as reported below.

handshape contour The signs for *disappear* and *favorite* are identified by handshape, point of articulation, orientation, and movement parameters. In *disappear*, the hands are open and face each

²The Analyzed Mexican words are reported in English for better comprehension.

³This means that the configuration of the hand at the beginning of the process of signing is different from the one at the end of the process



Figure 1: Gray and Lawyer. On the left side, gray, the hands performs a clockwise movement; on the right, hand moves from left to right to left.

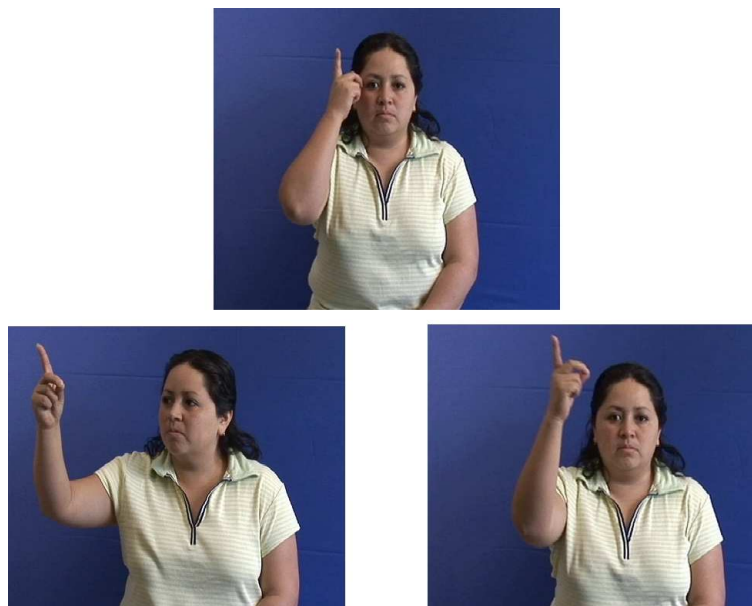


Figure 2: Owner and God. In owner, on the left, the hand is pointing to the front of the signer; for God, it points upwards.

other; the tips of all fingers of both hands make contact and then move away from each other. This is all the information we need to retrieve the sign. The fact that there is a handshape contour while hands separate does not have to be specified. On the contrary, for *favorite*, open hand configuration is selected, hand orientation is specified: the palm of the hand faces outward so it is the edge of the hand what makes contact with point of articulation (opposite shoulder) and then moves away from point of contact while changing aperture, as showed in figure 3;

handshape contrasts Contrary to handshape contours, the handshapes contrasts that we have found have to be specified for the signs to be

successfully retrieved. However, new lexicalized finger spelled forms for new objects (brands, places) or concepts may have multiple handshapes which would not allow the search to be as fast as desired.

We have proposed using the first and last handshape for the search for tagging videos. As reported in section 2.2.3., if videos are tagged ranking the sign parameters according to the information they provide without taking into account the handshape contour/contrast, the result will be that only one sign will be retrieved by the search engine.



Figure 3: In favorite, the hand changes hand-shape while signing.

2.2.6. Two-handed signs

Two-handed signs threatened to make the design of the search engine undesirably complex. The goal was to achieve the simplest representation of the non-dominant hand (H_2), the hand not usually used to articulate one-handed signs. Two-handed signs have been traditionally classified in three types, cfr. (Brentari, 1998). The list of two-handed MSL signs compiled were classified in two types only:

- *Type one* Signs in which both hands have the same configuration, perform the same motor act (synchronous or alternating) and do not contact each other. As none of these signs acquires a different meaning if articulated with just one hand, the sign can be retrieved just with the information of the dominant hand;
- *Type two* In all these signs hands make contact so non-dominant hand is considered (like other parts of the body) the place of articulation. Consequently, the information required for their retrieval is the same required by those one-handed signs presented in section 2.2.2.

2.3. Videos

The third part of the dictionary is the video manager section. Experts in both sign language and in educational curricula have selected the words that have to be included in the dictionary and “translated” into videos. The operation of “translating” words into videos is crucial for the dictionary. This part of the dictionary is shared between the two main sections described above and consists of a list of videos, tagged using the four sign parameters, which display Spanish words expressed in sign language. Each video is linked to the Spanish form and contains the sign parameters used to express that specific form. The tagging procedure is very important in the process of defining the bilingual dictionary. Videos are tagged

following a clear procedure which ranks the four sign parameters according to the information they provide to identify a sign.

The order of the parameters used to tag a video is the following: handshape, place of articulation, movement, orientation.

Place of articulation Place of articulation is used to differentiate two or more signs with same handshape. For example, the words *bitter* and for *captain* have the same handshape: five fingers (bent). But while for *bitter* the place of articulation is cheek for *captain* it is opposite shoulder of dominant hand, see figure 4;

Movement All monomorphemic signs contain a movement, but this parameter has to be taken into account only when the information given by the handshape and point of articulation do not lead to one specific sign;

Orientation Although the orientation parameter is considered the most difficult to describe, it does not present any serious challenge; it is only required by signs like *there is* and *to have*. They both show the same open hand configuration, make contact with the same point (middle of the chest) but in one sign the palm of the hand faces upward and in the other one it faces downward. (cf. *affection* in section 2.2.4.).

The biggest challenge we have to face is the complexity presented by lexicalized finger-spelled forms that show multiple handshape changes. We expect to encounter a number of this type of sign because bilingual schools are of recent creation in México, so most of the specialized words started as finger-spelled forms, but as mentioned in section 2.2.5. we have proposed using only two handshapes.



Figure 4: Bitter and Captain. Same hand-shape but different point of articulation.

2.4. The database

The structure of the database has an inner complexity because of the various purposes it has to satisfy. In fact, the database consists of three different schemes, so that each schema addresses one specific section of the dictionary as well as of the GUI.

The three schemes are described below:

catalog Analysis described above has allowed database designers to build a complex structured database which collects the four parameters in four dedicated tables. Such tables belong to this schema. Specifying one single table for each parameter has the benefit of creating a sort of registry of sign “alphabet”. The catalog schema contains tables to manage signs parameters as they are gathered and checked by sign language experts. However, this is not enough to manage subclassification of single parameters. Database designers have improved the catalog schema in order to manage subclassification of parameters by introducing some algorithmic tables;

video The video schema contains tables to manage external videos. Actual videos are saved in a dedicated file system, while this schema contains links and index for these videos. This schema controls links for both tagged and untagged videos. Video schema is the database counterpart of the middleware between MSL-Spanish and Spanish-MSL sections;

GUI The GUI schema is the most complex, since it manages tables for both the search engine and the administrative GUIs. For administrative purposes, we have also introduced a set of tables which provides facilities for loading and tagging videos.

3. Acknowledgement

Figures printed in this article have been extracted from videos recorded at the Servicios Audiovisuales -Universidad Autónoma Metropolitana (Iztapalapa). We would like to thank Dra. Lilia G. Ramos Santoyo and David Márquez for their precious help.

4. Final Remarks

This dictionary is the first Spanish-Mexican Sign Language / Mexican Sign Language Dictionary in México. The Spanish-MSL section of the dictionary, being a standard form enumerated dictionary based upon Department of Education grammar school textbooks, makes the basic education curricula accessible to Mexican deaf children. As their Spanish language competence has been proven to be lower than the average knowledge of the oral language of hearing impaired children in other countries, there is a serious need for this tool.

Most online dictionaries of well studied sign languages like British Sign Language⁴ (BSL), Aulan signbank⁵, American Sign Language Video Dictionary⁶, can only access the sign through the written word of the oral language. The Flemish Sign Language Dictionary⁷, a scholarly enterprise whose linguistics and sociolinguistics research was carried out by Ghent University and The Free University of Brussels allows signs to be read through the SignWriting System⁸, a transcription system that uses symbols to represent handshapes, movements, facial expressions ... After analyzing the SignWriting System (Sutton and DAC) we found out it was not suitable for the purposes of our dictionary.

The MSL-Spanish section of the Mexican bilingual

⁴<http://www.britishsignlanguage.com/>

⁵<http://www.auslan.org.au/>

⁶<http://www.aslpro.com/cgi-bin/aslpro/aslpro.cgi>

⁷<http://gebaren.ugent.be/information.php/>

⁸<http://www.signwriting.org/>

dictionary allows direct access to MSL signs and their Spanish translation through the browsing of the catalog of tagged videos. To create the alphabet necessary to access signs, we have analyzed the four traditional parameters of signed languages (handshape, place of articulation, movement, hand orientation). We call the attention to the fact that the presentation of Noema⁹ (electronic dictionary of Greek signs) only refers to handshapes as the tool that makes lemma searching possible.

The Graphical User Interface (GUI) provides an administrative interface designed to tag videos with forms and all four signs parameters. The e-Lis¹⁰ dictionary (Electronic Bilingual Dictionary of Italian Sign Language-Italian) is one of the few dictionaries that aims, like our bilingual dictionary at offering a search engine that allows the user to directly search for a sign, without resorting to the corresponding oral language term. The Multimedia Dictionary, (Wilcox, 2003), of American Sign Language designed an interface that allowed the user to search for signs by specifying for handshape, location and movement. Although the project began prior to the development of much that it is now recognized as multimedia technology and it never saw the light of day it raises issues that concern us:

How do we store and tag a signed language corpus?

Who constitute the usage panel?

The goal of the Mexican bilingual dictionary, and its intended users, have determined important aspects of the design of the search engine; the children have to easily identify the features they need (the four sign parameters) to retrieve the signs or words they need, and this task has to be carried on fast.

Consequently, our main concern has been creating alphabets, of the four parameters accessible to children's perception.

An important feature of the Mexican dictionary is that we have been able to simplify the searching engine, for example, by retrieving two-handed signs with the same mechanisms used to retrieve one-handed signs. The proposed dictionary will be finished within the bilateral agreement signed between the "Institute of Computational Linguistics" and the "Universidad Autónoma Metropolitana". This agreement will allow both participants to develop the dictionary and provide it to deaf communities as a value-added product with a strong social impact on such communities.

⁹<http://www.xanthi.ilsp.gr/noema/>

¹⁰http://elis.eurac.edu/index_en/

5. References

- Kai Alter and Angela D. Friederici. 2004. Lateralization of auditory language functions: A dynamic dual pathway model. *Brain and Language*, 89(2):267–276, May.
- Diane Brentari. 1990. *Theoretical Foundations of ASL Phonology*. Ph.D. thesis, University of Chicago.
- Diane Brentari. 1998. *A prosodic model of sign language phonology*. MIT Press.
- Angela D. Friederici. 2005. Neurophysiological markers of early language acquisition: from syllables to sentences. *Trends in Cognitive Sciences*, 9(10):481–488, October.
- Antoinette Hawayek and A Capistrán, 2008. In *Lingüística Formal*, chapter Deixis y marcación de persona en Lengua de señas mexicana (LSM). Universidad Autónoma Metropolitana-Iztapalapa, México.
- Antoinette Hawayek and Giuseppe Cappelli. 2004. Identificación y recuperacin del sujeto nulo: evidencia de una lengua signada (lsm). serie memorias. *VII Encuentro Internacional de Lingüística en el Noroeste.*, 1:411–430.
- Antoinette Hawayek. 1997. Las lenguas de seas en el diseo del lenguaje biológico. In *Memorias. Encuentro Nacional: Formación desde el contexto educativo, Guanajuato: Secretaría de Educación, Guanajuato*.
- Antoinette Hawayek. 2006a. La competencia lingüística de niños de 6^o año de dos escuelas de sordos. Technical report, CAM 39, IPPLIAP. Informe UAM.
- Antoinette Hawayek. 2006b. La contribución del espacio a la construcción del significado. *Lingüística Mexicana*, III(1):37–43.
- Antoinette Hawayek. 2007. La estructura del lenguaje a la luz de la adquisición y de las lenguas signadas. *Psicología y Evolución 3. Filosofía de la Ciencia, Enfoques Evolutivos y Metacognición. México: UNAM*.
- Geidy María Teresa Maya Anguiano. 2002. Enseñanza del español en niños sordos como segunda lengua a través de la lecto-escritura. Master's thesis, Universidad Autónoma Metropolitana-Iztapalapa, México.
- David Poeppel and Monahan Philip J. 2008. Speech perception: Cognitive foundations and cortical implementation. *Current Directions in Psychological Science*, 17(2):80–85, April.
- Philip M Prinz. 1988. Asl proficiency and english

- literacy acquisition: New perspectives. *Topics in Language Disorders.*, 18(4).
- Claire Ramsey. 1997. *Deaf children in public schools. Placement, Context and Consequences.* Gallaudet University Press.
- Regina Minerva Romero Alpizar. 2002. La competencia lingüística en español de un grupo de niños sordos en primero de secundaria. Master's thesis, Universidad Autónoma Metropolitana-Iztapalapa, México.
- Antonina Saba and Giuseppe Cappelli, 2003. *Morf-sin and AyDA: two systems for analyzing modern and old Spanish*, pages 865–890. *Linguistica Computazionale A Pisa.*
- Maricela Taba Nambo. 2003. Los niños sordos y la adquisición del español. Master's thesis, Universidad Autónoma Metropolitana-Iztapalapa, México.
- Sherman Wilcox. 2003. The multimedia dictionary of american sign language: Learning lessons about language. *Sign Language Studies*, 3(4):379–392.