Linguistic, Sociocultural and Technical Difficulties in the Development of a Spanish Sign Language (LSE Corpus)

Patricia Álvarez Sánchez, Inmaculada C. Baez Montero y Ana M. Fernández Soria

Construction and Exploitation of Sign Language Corpora

INITIAL WORK (1995)

Linguistic corpora have come to fill a privileged position because they constitute a valuable source of information for the creation of dictionaries, computational lexicons and grammars. [...] As a result, a new discipline appears: CORPES LINGÜÍSTICOS aimed at the processing and exploitation of this type of linguistic resource.

AIMS

Orienting LSE textual corpora of Galician signs from which to study the research on LSE.
A) Strengthening the corpus of LSE.
B) Determining which are the relevant linguistic units in SL.
C) Knowing the grammatical relational processes.
D) Developing tools for research: labelling, transcription, etc.
E) How to create a Corpus?

1. Create corpus real data
2. It must contain an irreducible base for linguistic description
3. It must be completed with computing support in order to make it easy to use.
4. It must include information of data
5. Different types of discourse samples
6. Some aspects typically, the type of discourse we want to obtain, etc.

PROCESS STAGES

1. Tool design for the creation of a corpus
2. Criteria for the selection of informants
3. Creation of a database of informant details
4. Collection of language samples
5. Data storage
6. Data labelling and marking
7. Transcription and notation systems

DIFFICULTIES

1. The lack of a research tradition on Sign Language in Corpus Linguistics forces us to seek paths from the very beginning
   - How to define units in sign languages
   - How to label the different semiotic elements for their later analysis
   - Other related issues
2. Creation of social networks in the Deaf community and the need to consider the social identity of our informants to be discussed.
3. Technical restrictions. We have to select appropriate material in order to avoid problems in the database, between the different digital video cameras, video player, computer, software,...

The study of LSE should not deal with a different manner to that of any other oral language.

It will be mandatory to have a textual corpus.

The production of a sign language has a intrinsic nature. Its reception is visual, so the conventions in sign language have to be recorded in video format.

ANALYSIS AND RECONSIDERATIONS

If our research answers to maintain essential or unessential information, we have made a good service to linguistics; however, the type of study usually depends on certain knowledge and experiences that do not correspond with the young researchers. (Lopes Morales, 1994, 29)

WHERE DO WE COLLECT OUR DATA?

We have developed an extensive field work with the purpose of representing the social and linguistic profile of the deaf community of people that have never registered in any textual corpus.

This is the data gathered from our SL informants:

1. Allocation: name, address and phone number (for contact)
2. Social and social environment (place and date of birth, age of deafness occurrence, deafness degree, deaf related family, job of a related family members)
3. School degree and type of studies, social situation, school situation of SL acquired.
4. Linguistic skills in LSE, in oral Spanish, in écrit, in English, etc.
5. Place of residence: in order to reflect and control linguistic curation.

DISTRIBUTION BY GENERIC TYPES

Gender and age
- Gender and sex: 25 minutes
- Age and sex: 23 minutes

Singers are interviewed on several topics, depending on their age, sex, preferences, etc. Thus, the discourse is more spontaneous.

Public discourse: 15 minutes

Correspondence and social talks are of a more formal and fixed style.

DISTRIBUTION OF INFORMANTS BY AGE GROUPS

- From 18 to 30 years: 35
- From 35 to 60 years: 5
- Over 60 years: 3
- Total: 43 interviews

FOR THE TIME BEING

AVANTES

We are members of a network of researchers for the teaching and research on Spanish and Catalan Sign Language.

We participate in the creation of the LSE Standardisation Centre (Laure B 2007).

We have obtained a grant (I.14.1.1) for our research project “Methods for the linguistic analysis of the Deaf Community”.

We consist of deaf teachers and researchers for the research and teaching work. We also consist of specialized researchers in deafness that will be the deaf teachers and teaching aids in the corpus.

We have achieved several works and publications of the PhD students in technical language. We have published. Other members of this group have published research papers on grammatical aspects in doctoral works.

ADVICE

Creating a textual corpus of LSE is a basis for:

1. Development of LSE grammars. The grammatical analysis will focus on the determination of the specific LSE rules and the grammatical processes in relation.
2. Applied research:
   - LSE interpretation
   - LSE codification
   - Normalisation and linguistics planning
   - Ancestral research
3. General research
4. Linguistic discourse
5. Linguistic investigation

OTHER RELATED ISSUES

4. Use of the corpus in the teaching platforms as a didactic element in order to provide the pupils with real linguistic situations. This will complete the kinesthetic learning process started within the classroom.

LINGUISTIC AND SOCIOLOGICAL CULTURAL DECISIONS

LSE is not an standardised language and there are very few descriptive studies on this language. This forces us to propose what kind of recording do we need, how many people do we need to be in the corpus to be representative and sound, and, finally, what concrete methods we would obtain from.

Taking into account these determining factors, we name:

- Asking for the cooperation of signers of different regions of Spain for obtaining a good representation of the different geographical regions.
- Selecting signers that fulfill certain features: native signers of LSE, professional signers of LSE and interpreters.
- Interview design:

Choice of deaf interviewees. Their database are more natural and they obtain a higher degree of involvement from the Deaf community in the project.

Records should be adapted to the presence of the interviewers.

We have prepared several models of the interview, with questions that may assess informants in the thematic fields, deafness, family, friends, relationships, thesaurus, etc.

TECHNICAL DECISIONS WITH A VIEW TO THE FUTURE

1. Standardisation of the recording format.
Use of a recording we digital cameras, similar work background in all the recordings, identical light conditions, signers clothes, positions and facing.

2. Multiple views of the signer: face, trunk, in profile.

3. Storage and handling of the recordings from the cause to the computer.

4. Editing of the recordings in chapters transcriptions, transcribed interviews and linear transcription for a better handling of the language.

5. Use of the ELAN system for the annotation process.

6. Computer cutting of grammatical features and sign configuration.

7. The use of gloss to annotate the modifications between interviews or research groups with the aim of determining our hand for future international use or work teams and on the other hand, the integration of results.

8. Eligible search and retrieval by sign configuration, grammatical possibilities, etc.

9. Online publishing of the corpus with the aim of external referencing.

BIBLIOGRAPHIC REFERENCES


Referred to the book "Sign Language", 1994, Editions Lenguaje Acta,

Slavonic letters

Versión 1  Versión 2

View

He comprado un nuevo juego

I have bought a new game

Observaciones

Fecha de entrada: 19/4/2004

Fecha de salida: 19/4/2004

Você pode querer uma nova vida