An Interdisciplinary Approach to Sign Language Data Set Collection of Native Signers

Miguel Rodríguez-Mondoñedo
Pontificia Universidad Católica del Perú
Lima, Peru
mmondon@pucp.pe

Francisco Cerna-Herrera
Pontificia Universidad Católica del Perú
Lima, Peru
francisco.cerna@pucp.edu.pe

César Ramos
Pontificia Universidad Católica del Perú
Lima, Peru
cesar.ramosc@pucp.edu.pe

Alexandra Arnaiz
Pontificia Universidad Católica del Perú
Lima, Peru
alexandra.arnaiz@pucp.edu.pe

Ana Rosa Moreno
Universidad Nacional Mayor de San Marcos
Lima, Peru
anarosa.moreno@unmsm.edu.pe

Joe Huamani-Malca
Pontificia Universidad Católica del Perú
Lima, Peru
huamani.jn@pucp.edu.pe

Gissella Bejarano
Marist University
Poughkeepsie, USA
gissella.bejarano@marist.edu

Abstract

Our work presents the design of a process to collect a sign language corpus from Deaf participants as an interdisciplinary effort. The process was designed to provide a guide to record videos ensuring that the linguistic data accurately reflects the natural usage of signers. A key aspect of this research was the direct participation of Deaf individuals as primary language informants and co-designers, reinforcing the authenticity and cultural relevance of the collected data. Moreover, an interdisciplinary team composed by linguists, interpreters, and engineers tested and refined the data collection methodology. The proposed process was designed to align with the linguistic and communicative norms of the Peruvian Deaf community. The data was gathered through signing tasks piloted with two Deaf participants. This process represents a significant step toward documenting and analyzing an LSP corpus. It provides a valuable resource for data collection involving minority groups, linguistic research, language planning, and promoting accessibility for the Peruvian Deaf community. The study underscores the importance of community-centered approaches and interdisciplinary collaboration in sign language research.

CCS Concepts

• **Human-centered computing** → Accessibility; Empirical studies in accessibility; Accessibility; Collaborative and social computing; • **Computing methodologies** → Artificial intelligence; Natural language processing.

PETRA '25, Corfu Island, Greece

ACM ISBN 979-8-4007-1402-3/2025/06
https://doi.org/10.1145/3733155.3734907

Keywords

sign language recognition, iconicity, Peruvian Sign Language

ACM Reference Format:
Miguel Rodríguez-Mondoñedo, Francisco Cerna-Herrera, César Ramos, Alexandra Arnaiz, Ana Rosa Moreno, Joe Huamani-Malca, and Gissella Bejarano. 2025. An Interdisciplinary Approach to Sign Language Data Set Collection of Native Signers. In *The PErvasive Technologies Related to Assistive Environments (PETRA '25), June 25–27, 2025, Corfu Island, Greece.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3733155.3734907

1 INTRODUCTION

The goal of this paper is to show how we can build an instrument to collect data from the Deaf community to train an AI model that recognizes the iconic relation between a sign in Peruvian Sign Language (LSP, its Spanish acronym) and the corresponding object it refers to (i.e., its referent). We describe the design and testing of a pilot in which Deaf Peruvians participate at each step. We contend that including this requirement in the development of any technological project profoundly deepens its positive impact on the target community, truly contributing to their empowerment. Technology is not just an exercise in high-tech skill, but the orderly transformation of our environment to improve the lives of people; to do that, we must understand their actual needs, not just imagine them or presuppose them. For that reason, the final users must be extensively consulted and their input collected and analyzed at every stage of the process. These requirements become even more urgent when the goal is to develop technology for people with special needs, as with the Deaf community.

The paper is organized into five sections. The first one is this introduction. In the second one, we briefly describe the AI model for which we are building and piloting a data collection instrument. The third section presents the properties of the pilot and the steps taken to develop it with Deaf participants. The fourth one discusses the findings from this endeavor. The final section offers the conclusions.

2 A MODEL TO RECOGNIZE ICONICITY

As pointed out by several researchers, AI still needs to find strong fundamentals to evaluate causality and do reasoning. This work focuses on understanding iconicity, a special case of analogy [1], as an innovative research direction to contribute to reasoning in AI. More specifically, we focus on the process of collecting a corpus that is used to build an AI model that can learn to detect iconicity or the implicit relation between some signs and objects/actions and extend that learning to new signs.

2.1 Searching for iconicity in LSP

We define iconicity in the classical way, going back to Peirce [2], as a relation of similarity between a sign and the object it refers to. While well established, this definition presents several complications [3–5], and no consensus has been reached. The definition has faced multiple theoretical and methodological challenges, as researchers must decide from which perspective or approach to tackle it [6]. One can take a perspective of iconicity based on the intuition of the researchers or the participants, or a one based on the effective recognition of iconicity by users of a particular semiotic system. This is relevant because the aims, nature, and possibilities of the research will determine the choice of one approach over another, which will naturally also influence the results. The problem with the definition lies on the idea of "similarity", which is not a categorical variable but an interval one, meaning that iconicity is better measured with a scale. Furthermore, iconicity is the result of a representational process [2], i.e., it is a relation linking three elements: a material form (gesture, sound, etc.), a referent (a particular segment of reality) and an interpreter (an agent to whom this relation holds). In other words, a sign is iconic if and only if its form is recognized as similar to some referent by an interpreting agent. If we combine its scalar nature with its need for an interpreter, we conclude that iconicity may be different to different agents in various degrees. Thus, since our goal is to build a computational model that recognizes iconicity, that is, an agent that interprets the relation between a sign and a referent as iconic (i.e., that establishes what we will call an iconicity-based sign-referent pair), we need to build an instrument to collect data taking all the above properties into account.

For these reasons, the key question we need to answer to build the proper instrument for this task is the following: How do we determine the degree of iconicity of a given sign? Given the crucial role interpretation plays in the definition of iconicity, the only possible answer is by asking the relevant interpreting agents. In the current case, this means asking Deaf consultants natively proficient in Peruvian Sign Language (LSP). The computational model must learn from the judgments of these individuals to interpret other pairs of sign-referents as iconic.

LSP is a full-fledged language, created by the Deaf community in Peru (for a basic LSP grammar see [7]). It uses movements of the hands, face, and body in the space to construct lexical items, which are expressed in the visual modality. By using handshapes and movement as a tool to produce meaning, it has the possibility to iconize object shapes and contours (which is different from oral languages, which can only iconize with sounds). Given that all visible objects have shapes and contours, this predicts a much higher amount of iconicity in sign languages than in oral languages, as it is the case [8]. Sign languages are, therefore, well suited to feed systematic data to computer models that could act as agents able to recognize iconic relations.

2.2 Model design and the data set

The model for which we designed this data collection process aims to understand sign language, gestures, and the role of iconicity in Deaf culture. It proposes a new and innovative direction for exploring analogy as a type of reasoning. In deep learning, models sometimes produce correct outputs for the wrong reasons, so improving their robustness is essential for building more trustworthy systems. This AI model can also help draw attention to a more immediate challenge: developing sign language translation tools and bridging the gap in access to information and education for Deaf communities. This issue is especially pressing in developing countries like Peru, where very few Deaf individuals have access to higher education.

To measure gesture and sign language understanding, we need to define a list of highly probable, culture-independent signs (essentially, a common denominator across cultures), which improves on previous attempts [9–11]. However, proposing such a list requires culturally specific knowledge to ensure the selected signs are appropriate for analysis. The recording task is not straightforward, as it requires prior knowledge of both the Deaf community and familiarity with sign language. This is especially important, as specific linguistic discussions may arise during the recording sessions, such as questions about a sign's status or meaning. For example, a signer may produce a personal, spontaneous sign that is not (yet) part of language's grammar and should not be considered as though it were. Another example concerns the semantic relationship between signs: to what extent can a similar sign be considered a variant of the same word, or should it be considered another word? The very definition of what constitutes a 'sign' (or 'word') remains a matter of debate and requires clear criteria. Prior experience is also important for ensuring high video quality, which is crucial for subsequent analysis. This includes knowing where to position the camera(s) and how to frame the signer, keeping in mind the needs of the later stages of analysis. After recording, the raw video data must be manually processed to isolate the specific signs for analysis. For this step, we will collect video clips of signs paired with images representing the corresponding object or action. The signs will then be grouped into those known to exhibit iconicity and those that do not.

To develop an AI capable of recognizing new signs based on pre-identified iconicity properties, we will train a generative model (e.g., StyleGAN [12]) to convert sign language videos into still images, using only training signs. These output images will serve as input to a pre-trained image recognition classifier, which will be fine-tuned using the images generated by the model. If the trained AI system can successfully classify the gestures, this indicates that it is learning relevant patterns for recognizing new, unseen signs.

2.3 Method limitations in other projects

As mentioned, the main purpose of this paper is to show how valuable the contribution of Deaf participants is in developing tools

Table 1: Comparing different datasets with respect to the participation of Deaf people

Name	Country	Participants	Authors affiliations
DGS Kinect 40 [20]	Germany	15 signers (it does not say if Deaf)	Centre for Vision Speech and Signal Processing - University of Surrey (all engineers or scientists)
RWTH-PHOENIX-Weather [21]	Germany	9 interpreters	Human Language Technology and Pattern Recognition, RWTH Aachen University, Germany
SIGNUM [22]	Germany	25 native signers	Human Language Technology and Pattern Recognition, RWTH Aachen University, Germany
GSL 20 [23]	Greek	7 GSL signers	The Visual Computing Lab Information Technologies Institute Centre for Research and Technology Hellas
DEVISIGN-G [24]	China	8 signers	Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology
WLASL [25]	EEUU	119 from ASLU, ASL-LEX and ASL Youtube videos (no guarantee if they are native or non native)	The Australian National University, Australian Centre for Robotic Vision
CSL-Daily Dataset [26]	China	10 deaf signers, 4 sign teachers	Institute of Artificial Intelligence Hefei Comprehensive National Science Center, China
NMFs-CSL Dataset [27]	China	10 signers	Institute of Artificial Intelligence Hefei Comprehensive National Science Center, China
BSL [28]	Britain	237 signers who reported learning to sign before the age of 7, 12 signers after 7 years old	La Trobe University College London
AUTSL [29]	Turkish	43 signers (among then only 1 deaf, 1 CODA)	Computer Engineering Department Ankara University
LSA64 [30]	Argentina	10 non-expert signers	Instituto de Investigación en Informática LIDI, Universidad Nacional de La Plata

for collecting data to train iconicity-based AI recognition models. Naturally, we aim to provide our models with reliable data. Sign languages vary across countries and even between communities. These languages have native signers, individuals who have grown up within Deaf communities [8], and who can produce signs that reflect socially recognized usage. It is therefore questionable to collect signs from other types of signers, such as interpreters or hearing learners of sign language as a second language. Several sign language datasets have been developed in different countries, each with specific design choices regarding participants and data collection. Previous research has provided valuable resources, but often lacks input from Deaf signers or does not specify signer profiles in detail. We contend that this is a serious limitation that hinders progress in this area of research. Table 1 summarizes the information of some of these datasets, highlighting how participant selection varies and often excludes native Deaf signers.

In that sense, our project aims to improve how data is collected for technological initiatives, especially those aimed at improving the lives of people with disabilities.

3 DESIGNING THE PILOT

We conceived an instrument consisting of 100 LSP signs to be presented to 50 Deaf signers, who must rate them on a 5-point Likert scale. As mentioned, this paper describes the process of building and testing the pilot. To build it, the first stage was to select 100 signs, and the second was to determine the best way to elicit the ratings. Each step is described in the following subsections, with a focus on cooperation with the Deaf community.

3.1 Selecting the signs

The selection was carried out by a team comprising one Deaf consultant, one LSP interpreter, three linguists, and two engineers. We decided to focus on signs with direct visual ties to their referents, whether entities, actions, or states, and this became the main inclusion criterion for the list. This is because iconicity-based sign-referent pairs can exhibit other types of visual associations. For instance, the sign NOSE (here, we follow the convention of using capital letters to refer to signs) is produced by pointing to the nose. It is also possible for a sign to represent an abstract concept by iconizing a culturally associated visual object. For example, JUSTICE, which is produced by moving the hands to simulate a weighing scale. There are also signs that change form depending on their structural position in a sentence. For instance, verbs may take on a specific handshape (called a classifier [8]) attached to them in 'The book fell from the shelves' versus 'He put the book on the shelves,' even though in both cases the handshape iconically represents the book.

Consequently, we excluded these types of signs: body parts (because they are typically referred to by pointing to the corresponding part on one's body), abstract nouns and verbs (due to their reliance on cultural iconicity), and classifiers (as they are grammatical elements whose properties depend on syntactic distribution). We also excluded signs requiring specific cultural knowledge to be understood, as they are too region-specific. For example, the sign for ANTICUCHO, a popular Peruvian dish made with skewered and grilled meat, typically beef heart, was excluded, since we intend the list to be applicable to other sign languages. Lastly, initialized signs were also excluded. Initialization is a productive mechanism in several sign languages, whereby a sign is formed using one or more letters from the manual alphabet of the sign language, which is typically associated with the dominant oral language of the region. For example, in LSP, one sign for PAPAYA (also papaya in Spanish) is produced by performing the three letters from the LSP manual alphabet: P, P, and Y. While initialized signs may exhibit a form of iconicity, it does not operate directly between the sign and its referent. Rather, the relationship is mediated by a third element: the written form of the word in the oral language. In this sense, the sign for papaya, composed of the letters P, P, and Y, is not iconic of the fruit itself, but rather of the orthographic form of the word 'papaya'.

3.2 The elicitation task

Once the selection of 100 signs was completed, the next step was to determine the best way to elicit ratings from the group of 50 Deaf signers. Three key issues had to be considered. First, we needed to decide on the most appropriate channel for presenting the stimuli: whether through written words, pictures or videos of the intended referents, or LSP in a repetition task. Second, it was essential to clearly explain the goals of the project, particularly the concept of iconicity. Third, we needed to obtain informed consent from the participants.

The literature discourages the use of orthographic words when collecting data from Deaf signers, as establishing a one-to-one equivalence between a written word and a sign is challenging and often results in diverse or unintended outcomes [13, 14]. In our data collection task, presenting an orthographic word to a Deaf participant could lead to the production of signs different or unrelated to those intended in the list. Moreover, sign language research is best conducted using a monolingual approach; introducing another language and its writing system may cause unintended transfer of linguistic features. This is especially important in Deaf communities where access to literacy education has been limited. In such contexts, research should rely on images or visual stimuli instead of written words, which may not be fully understood or misinterpreted [15]. For these reasons, we decided against using Spanish text and instead opted for visual stimuli.

Visual stimuli are the preferred approach when studying sign languages; however, the effectiveness of data collection depends on how well these stimuli align with the research goals. For instance, signers are more likely to associate a sign with an image when there is a clear and logical correspondence between them: for example, the sign for BASKET in American Sign Language is more easily matched to an image of a round basket than a square one [16]. Using pictures or videos is also appropriate when working with signers who have limited proficiency in written language. Nevertheless, this approach can generate a wide range of responses. A single image or video may evoke not only multiple interpretations but also different levels of linguistic expression, from a single word to full sentences. Even when a signer responds with a single sign, it may not correspond to one of the target signs from the intended list [14]. In short, while visual stimuli are essential for accessibility and iconic mapping, they can also introduce variability that challenges the specificity required for our research objectives.

For these reasons, we conclude that the most effective method for documenting the list of signs from Deaf users is to present them with a video of each sign, previously recorded by another Deaf signer, and then ask them to reproduce that specific sign. Initially, we considered asking participants to also reproduce any alternative signs that conveyed the same meaning as the one shown, but ultimately, we decided against this due to time constraints. This approach constitutes a repetition task, enabling us to capture the intended signs from the list with greater accuracy. A similar methodology was adopted in [17], where 309 signs were recorded following a selection process, and ratings were obtained from a group of 20 Deaf individuals.

3.2.1 Designing the tasks. It must be considered that most Peruvian Deaf people have not received adequate literacy education [18, 19], resulting in a significant proportion of them exhibiting limited proficiency in written Spanish. For that reason, it would have been detrimental to our goal (besides being culturally disrespectful) to provide the explanation of the project, the instructions to do the rating, and even the informed consent, exclusively in written form. We also decided to present these materials as pre-recorded videos with explanations composed in LSP by Deaf signers.

We also decided to explain the notion of iconicity to the participants. For this purpose, a Deaf consultant (and co-author) was familiarized with this concept. She is not a linguist, so she had to be trained in the relevant concepts to become familiar with the precise characterization of iconicity, its fundamental problems, appropriate examples, and the specific goals of our project. After that, she participated in the selection process (described in previous sections) and was ready to compose an explanation suitable to be presented to the Deaf participants. To do this, she had to record a video in LSP, adapting the discourse to the educational background of the participants.

One question is why we decided to introduce the notion of iconicity to the participants, instead of building a simpler set of instructions around the notion of "similar to", thus avoiding a technical term that was completely new to all of them. We must keep in mind that almost no Deaf Peruvian has ever had any kind of training on any aspect of LSP grammatical, functional, or structural properties. It is certainly not part of any school curriculum (mainstream or otherwise). However, we believe that it is our duty as LSP researchers to introduce participants to the basic concepts on which our study is based, in this case, iconicity. This serves two purposes. On one hand, we are being transparent about our goals and the ways we will achieve them; on the other hand, we are encouraging our participants to think about their language by transferring a piece of linguistic metalanguage. In addition, the fact that a Deaf person, as part of the research team, delivers the explanation serves as evidence that LSP is perfectly capable of including technical concepts and that it is not a barrier for scientific communication.

Figure 1: In Pilot A (left), participants provided iconicity ratings while recording their signs; in Pilot B (right), iconicity ratings were submitted separately via Google Forms.

The last point may seem obvious to any informed person, but most people, including Deaf Peruvians, find this basic fact about LSP surprising. In this sense, we believe that we are contributing to the epistemic empowerment of the Deaf community. To achieve this goal, a Deaf co-author recorded a video in LSP explaining the notion of iconicity and the different degrees of it, with relevant examples. The vast majority have not participated in projects. The explanation is a teaching moment for the Deaf community. We are explaining how to do research in the academic domain. The value of their work, and their rights to their images, and always knowing where their data is. Because the one explaining is Deaf, the project has more value. This can be used as a precedent for other projects, so that they can know their rights and that there is an ethics behind any serious research. In addition, because no one knew what iconicity was. We also had to explain the nature and goals of the project and ask for informed consent. For all of this, we decided to record the stimulus videos in LSP.

3.2.2 Testing two methods to collect the ratings. The videos were prepared by one Deaf consultant, under the supervision of an LSP interpreter (who is also a linguist). They consist of (a) informed consent, (b) an explanation of the concept of iconicity, (c) an explanation of the project's goals.

To determine what was the best way to deliver the instructions and to collect the ratings, we prepared a pilot study in two parts (A and B), dividing the input into two sets of 50 signs each. Each set had different methods to ask for the ratings. We piloted both sets with two different consultants at different times. We decided to give the instructions and explanations on video because that way everyone would have the same information, and we would have control over the information, and we could measure the time more appropriately. We also decided to play the videos at a 0.9x speed, so that everything was clearer, but, for the first consultant, this was not enough, so we switched to 0.7x, which worked better. Two sessions were conducted to estimate the time needed. At the beginning of each session, a video recording of the informed consent in LSP was played. Each session was designed to include the elicitation of 100 signs by the participants, with one sentence per sign, and a personal rating on a scale of 1 to 5 to assess the level of iconicity of each sign. In addition, participants were asked to create a sentence using each sign and to provide a variant of the sign, if one existed.

Two methods were tested, referred to as Pilot A and Pilot B. The main difference between them was how the iconicity ratings were collected: in Pilot A, the ratings were collected during the recording of the signs from the list, while in Pilot B, the ratings were collected using a Google Form, as illustrated in Figure 1.

In the Google Form method (Pilot B), the videos of the signs from the list had to be watched twice: initially for the Google Form and subsequently during the recording. Table 2 summarizes the differences between the two methods.

For each target sign, participants were shown a video containing step-by-step instructions to replicate the sign, give the iconicity rating, produce a sentence with the sign, and provide a variant. In the case of Pilot A, the signer was first shown a specific sign and then requested to reproduce it. Subsequently, they were asked to rate the iconicity of the target sign on a scale ranging from 1 to 5 using LSP numbers, which are recorded in video and registered by the researchers in an Excel sheet.

The second method included two materials: a video and a Google Forms survey. The survey incorporated videos of the target signs to facilitate the assessment of each sign's iconicity on a scale ranging from 1 to 5. After completing the survey in the Google form, each target sign was shown again, and for each one, participants were shown a video with step-by-step instructions to replicate the sign, produce a sentence, and provide a variant if one existed. In other words, the rating was done first, and then the rest of the stimuli were shown.

The purpose of the pilot study was to compare the two methods in order to get an initial idea of which one required less time and which was more user-friendly for Deaf individuals. Although we did not conduct a statistical comparison, the sessions provided insights for this evaluation. Sessions were conducted with two participants of different ages, one under 30 years old and the other over 40, to explore whether age influenced the speed of one method compared to the other.

3.3 The videos

To implement the methods, a series of videos were recorded in advance with a Deaf individual who is part of the research team and a coauthor of this paper. The video featured a set of 100 previously selected signs. Each segment included two main components. First, it presented the target sign along with its corresponding word in Spanish, as shown in Figure 2 left. Then, the instruction COPIAR SEÑA ('copy sign') appeared on screen as shown in Figure 2 right, while the target sign was shown again to allow the participant to replicate it.

Next, a black screen displayed the text "NIVEL DE ICONICIDAD" ('iconicity rating'), accompanied by a simple graph illustrating a 5-point Likert scale using dashes (1—————-5), prompting the participant to rate the degree of iconicity of the target sign using LSP.

Table 2: Comparing the two methods

Pilot A (50 signs)	Pilot B (50 signs)
One continuous activity: recording (iconicity ranking asked during recording)	Two activities: (a) iconicity ranking in Google Forms, (b) recording
Sign videos are shown only once during the session	Each sign video is shown twice during the session: once for completing the Google Form and once for recording the sign
Participants are asked to make a sentence and give a variant sign (if there is one)	Participants are asked to make a sentence and give a variant sign (if there is one)

Figure 2: The sign and its repetition with the instruction to copy the sign

Figure 3: In both pilots, participants copied the sign they observed, and we tested the use of a sheet of paper placed in front of the camera to mark the beginning and end of each recorded sign, allowing automatic segmentation of the videos afterwards.

After that, another screen appeared with the text "ORACIÓN" ('sentence'), inviting the participant to create a sentence incorporating the target sign. Finally, the text "VARIANTE" ('variant') was shown, allowing the participant to offer an alternative sign if they knew one.

In both pilots, we tested the use of a sheet of paper as a visual marker during the recording to indicate when a participant began and ended their signing, as can be seen in Figure 3.

To separate each piece of data, a designated team member was responsible for placing the paper in front of the camera at the appropriate moments. Another method tested for marking these transitions involved instructing participants to lower their hands after completing each signing task. The purpose of this strategy was to facilitate automated video segmentation: by programming a script to detect either the appearance of the paper or the resting position of the hands, we could identify precise cutting points for the video clips.

4 DECISIONS MADE BASE ON THE PILOT'S RESULTS

In this section we present the decisions we made after analyzing the pilot, in order to build the final instrument to collect our data.

4.1 Preliminary meetings via Zoom

It was observed that the consultants in the in-person session appeared fatigued and unengaged after the video explanations. Consequently, in collaboration with the consultants, it was decided that the initial meeting would be via Zoom, with the consultants to show the instructional videos beforehand. This allows for additional time for questions. This does not preclude presenting the informed consent and instructions during the actual session. Since the topic is new to them, it is better to follow a protocol where the Zoom meeting is led by a Deaf person to promote mutual trust. Deaf people must be involved in the discussion of the entire process, from the selection of signs to the pilot, including videos, instructions, examples, and, of course, ratings. During each Zoom meeting, the project was explained, the concept of iconicity was introduced, and a translation of the informed consent was shown; each meeting lasted approximately an hour and a half. In these meetings, potential candidates asked any questions they might have, and those who were interested ultimately shared their contact information with the research team. These meetings facilitated recruitment while also providing an opportunity to present the project and initial concepts, which helped to save time during the sessions.

4.2 The method for marking video cuts is lowering hands

We tested methods for marking video cuts, either by placing a sheet of paper in front of the camera or by having participants lower their hands to the default position. Ultimately, we chose the second option, since using a sheet of paper required a dedicated person during recording and was prone to errors. The instruction to lower their hands is simple, and with some practice, participants become accustomed to doing so at the end of signing.

4.3 Iconicity rankings collected via Google Forms

Faced with the option of recording the sign, sentence, and iconicity ranking consecutively or collecting the iconicity ranking through a Google Form, we ultimately decided on the latter. This approach lets participants concentrate on copying the sign and creating the sentence during the recording, and focus on the rankings separately. Additionally, by asking them to complete the Google Form before the recording, participants are exposed to the signs they will later need to copy and include in their sentences, as these are the same signs. This familiarity makes the recording process smoother.

4.4 Variants not collected

Although we initially considered collecting sign variants, it was ultimately discarded. This was because providing variants became an additional task during the session, making the recording process more demanding for the participants. Furthermore, it was observed during the pilots that not all signs on the list had a variant, and the most common scenario was the absence of variants altogether.

4.5 Spatial organization

During the pilot sessions, a spatial organization was tested in which two researchers were positioned near the participant to answer any questions and discuss the concept of iconicity with them. Additionally, one researcher was responsible for operating the camera, while another observed to ensure the recording process flowed smoothly. For the pilot sessions, two cameras were utilized: one positioned in front of the participant to record the data and another to capture the entire session. For the actual recordings, only the camera in front of the participant is needed, along with a minimum of two researchers: one to manage the camera and another to interact with the participant. Furthermore, it was decided to place a monitor next to the camera, with the objective of displaying the signs from the designated list. Figure 4 shows this spatial organization.

4.6 The structure of the session

We decide that the session should proceed in the following sequence. First, the participant is provided with the informed consent

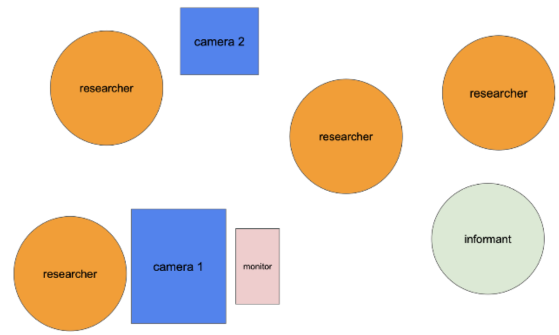

Figure 4: Camera settings during the pilot.

document to sign and agree to collaborate on the research project. This document had previously been explained in LSP during the virtual meeting mentioned in 4.1. Still, if the participant wants to watch the informed consent video in LSP, they could request it. They could also ask any questions they needed before signing the document. Once the consent form has been signed, a video is shown explaining the concept of iconicity in LSP, followed by the presentation of a video that describes the various levels of iconicity (ranging from 1, least iconic, to 5, most iconic) that the participant should use to categorize the signs they will observe during the session. Subsequent to the presentation of both videos, the participant is asked about their understanding of the concept of iconicity. They are also asked to provide their examples and to evaluate specific signs. This procedure is used to ensure that the participant has comprehended both the concept of iconicity and the task they will be undertaking in the research. After this introduction, a video is presented that provides a detailed explanation of the task to be performed (the ranking of the 100 signs on the list from 1 to 5 using a Google Forms survey). At this moment, the participant is provided with a laptop to complete this task. Upon completion of the form, two additional videos are displayed. The first video provides a detailed explanation of the subsequent task, which is to record the signs. The second video outlines the requirement to produce a brief sentence following the recording of each sign. Once the participant has viewed both videos and has no questions, we proceed to record the 100 signs along with their corresponding sentences. The recording will be made with no interruptions, and the segments will be automatically divided based on the lowering of the hands, which will serve as a cue.

4.7 The structure of the videos

Based on the lessons learned from the pilots, we created the final video using the signs from the list. This video is about 20 minutes long, contains 100 signs, and is structured as follows. First, each sign is shown individually so that the student can become familiar with it. Above the video of the sign, the Spanish word corresponding to the sign (CASA, GATO, ESTRELLA, etc.) is displayed in white capital letters on a black background. Next, the individual sign is shown again, but this time the instruction "repetir seña" ("repeat the sign") appears above the video, indicating that participants should copy the sign at that moment. The instruction "repetir seña"

Figure 5: The sign (left) and the instruction to repeat it (right).

Figure 6: The instruction to elicitate the corresponding sentence.

is written in white capital letters, but the background is red. Figure 5 shows the two steps for each sign.

The second and final task is to display the Spanish word "oración" ("sentence") in uppercase white letters centered on a black background, as illustrated in Figure 6. This indicates that participants must create a short sentence using the sign they have just seen and copied. A limited amount of time is allotted for this activity.

Accordingly, the video structure for each sign was as follows: the sign, the "repeat sign," and the "sentence." When employing the video as a tool, it is recommended that participants engage in a few preliminary trials to become acquainted with the structure of the video. Following this preliminary period, the video is to be played in its entirety. Participants are permitted to request a pause or to make mistakes. They may also request that the sign be repeated or inquire about its meaning if it is unclear. In such instances, the video and recording are stopped, and the video is replayed, commencing from the most recent instance of correct performance. This is how the sessions are conducted. The data obtained from the recordings will later be analyzed and processed. Additionally, as stated in the informed consent, the data from the recordings will be uploaded to the university's institutional repository. Keypoint landmarks extracted from the videos will be freely available, while the original (raw) videos will be accessible upon request for research purposes only.

5 CONCLUSIONS

A collection instrument for Peruvian Sign Language was presented with the objective of recording videos to train an artificial intelligence model. As has been demonstrated, a designed instrument for the collection of sign language data necessitates the involvement of members of the Deaf community not only as direct informants but also as consultants during the development of the appropriate instruments and the elicitation stimuli. The implementation of the instrument with Deaf participants enables the assessment of all the particular conditions associated with data collection, which enhances the probability of acquiring more valuable data. This is particularly evident in the context of AI models, where the quality of the data is of paramount importance.

Acknowledgments

This project was funded by the AI2050 Early Career Fellowship Grant G-23-66054 from Schmidt Science.

References

[1] Esa Itkonen. 1994. Iconicity, analogy, and universal grammar. Pragmatics.
[2] Charles Hartshorne, Paul Weiss, and Arthur Burks (Eds.). 1958. Collected Papers of Charles Sanders Peirce. Harvard University Press, Cambridge, MA.
[3] Juan Carlos Moreno Cabrera. 2020. Iconicity in Language: An Encyclopedic Dictionary. Cambridge Scholars Publishing.
[4] Susan Taub. 2001. Language from the Body: Iconicity and Metaphor in American Sign Language. Cambridge University Press.
[5] Marloes Oomen. 2021. Iconicity and Verb Agreement: A Corpus-Based Syntactic Analysis of German Sign Language. Walter de Gruyter.
[6] Yasamin Motamedi, Hannah Little, Alan Nielsen, and Julia Sulik. 2019. The iconicity toolbox: Empirical approaches to measuring iconicity. Language and Cognition 11, 2 (2019), 1–20.
[7] Miguel Rodríguez-Mondoñedo *et al.* 2023. Gramática Básica de la Lengua de Señas Peruana. Facultad de Letras y Ciencias Humanas, PUCP. Retrieved from https://www.youtube.com/playlist?list$=$PLTCzdajd5pCForZu712cwMYsLOz1pWaIS
[8] Wendy Sandler and Diane Lillo-Martin. 2006. Sign Language and Linguistic Universals. Cambridge University Press, Cambridge, UK.
[9] Shuangyu Yu, Carlo Geraci, and Naomi Abner. 2018. Sign languages and the online world: Online dictionaries and lexicostatistics. In Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC 2018).
[10] Morris Swadesh. 1971. The Origin and Diversification of Language. Edited posthumously by Joel Sherzer. Aldine, Chicago, IL.
[11] James Woodward. 2000. Sign languages and sign language families in Thailand and Vietnam. In Karen Emmorey *et al.* (Eds.), The Signs of Language Revisited: An Anthology to Honor Ursula Bellugi and Edward Klima. Lawrence Erlbaum Associates.
[12] Tero Karras, Samuli Laine, and Timo Aila. 2019. A style-based generator architecture for generative adversarial networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2019).
[13] Rachel Channon. 2015. Research methods for studying the form of signs. In Eleni Orfanidou, Bencie Woll, and Gary Morgan (Eds.), Research Methods in Sign Language Studies: A Practical Guide. Wiley-Blackwell, 123–140.
[14] Josep Quer, Carlo Cecchetto, and Caterina Donati. 2017. SignGram Blueprint: A Guide to Sign Language Grammar Writing. De Gruyter.
[15] Victoria Nyst. 2015. Sign language fieldwork. In Eleni Orfanidou, Bencie Woll, and Gary Morgan (Eds.), Research Methods in Sign Language Studies: A Practical Guide. Wiley-Blackwell, 108–122.
[16] Robin L. Thompson, David P. Vinson, and Gabriella Vigliocco. 2009. The link between form and meaning in American Sign Language: Lexical processing effects. Journal of Experimental Psychology: Learning, Memory, and Cognition 35, 2 (2009), 550–557.
[17] David P. Vinson, Kearsy Cormier, Tanya Denmark, Adam Schembri, and Gabriella Vigliocco. 2008. The British Sign Language (BSL) norms for age of acquisition, familiarity, and iconicity. Behavior Research Methods 40 (2008), 1079–1087.
[18] Elizabeth Parks and Jason Parks. 2010. A sociolinguistic profile of the Peruvian Deaf community. Sign Language Studies 10, 4 (2010), 409–441.
[19] Sara Goico. 2019. The impact of 'inclusive' education on the language of deaf youth in Iquitos, Peru. Sign Language Studies 19, 3 (2019), 348–372.
[20] Helen Cooper, Eng-Jon Ong, Nicolas Pugeault, and Richard Bowden. 2012. Sign Language Recognition Using Sub-Units. Journal of Machine Learning Research 13, 72 (2012), 2205–2231. http://jmlr.org/papers/v13/cooper12a.html

[21] Jens Forster, Christoph Schmidt, Oscar Koller, Martin Bellgardt, and Hermann Ney. 2014. Extensions of the Sign Language Recognition and Translation Corpus RWTH-PHOENIX-Weather. In Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC 2014), 1911–1916. https://aclanthology.org/L14-1472/

[22] Ulrich von Agris and Karl-Friedrich Kraiss. 2010. SIGNUM Database: Video Corpus for Signer-Independent Continuous Sign Language Recognition. In Proceedings of the LREC2010 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies, 243–246. https://www.sign-lang.uni-hamburg.de/lrec/pub/10006.pdf

[23] Nikolas Adaloglou, Theocharis Chatzis, Ilias Papastratis, Andreas Stergioulas, Georgios Th. Papadopoulos, Vassia Zacharopoulou, George J. Xydopoulos, Klimnis Atzakas, Dimitris Papazachariou, and Petros Daras. 2022. A Comprehensive Study on Deep Learning-Based Methods for Sign Language Recognition. IEEE Transactions on Multimedia 24, 1750–1762. https://doi.org/10.1109/tmm.2021.3070438

[24] Yuecong Min, Xiujuan Chai, and Hanjie Wang. 2014. DEVISIGN-G: A Chinese Sign Language Dataset for Isolated Sign Recognition. Visual Information Processing and Learning (VIPL), Institute of Computing Technology, Chinese Academy of Sciences. https://vipl.ict.ac.cn/homepage/ksl/data.html24

[25] Dongxu Li, Cristian Rodriguez Opazo, Xin Yu, and Hongdong Li. 2020. Word-level Deep Sign Language Recognition from Video: A New Large-scale Dataset and Methods Comparison. In Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV 2020). IEEE, Snowmass Village, CO, 1459–1469. https://arxiv.org/abs/1910.1100626

[26] Hao Zhou, Wengang Zhou, Weizhen Qi, Junfu Pu, and Houqiang Li. 2021. Improving Sign Language Translation with Monolingual Data by Sign Back-Translation. In Proceedings of the IEEE/CVF International Conference on Computer Vision and Pattern Recognition (CVPR 2021). IEEE, Nashville, TN, 1073–1082. https://arxiv.org/abs/2105.1239728

[27] Hezhen Hu, Wengang Zhou, Junfu Pu, and Houqiang Li. 2021. Global-Local Enhancement Network for NMF-Aware Sign Language Recognition. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV 2021). IEEE, Montreal, Canada, 1073–1082. https://arxiv.org/abs/2008.1042830

[28] Schembri, Adam, Jordan Fenlon, Ramas Rentelis, and Kearsy Cormier. 2011. British Sign Language Corpus Project: A corpus of digital video data of British Sign Language 2008–2011 (First Edition). University College London. https://www.bslcorpusproject.org29

[29] Ozge Mercanoglu Sincan and Hacer Yalim Keles. 2020. AUTSL: A Large Scale Multi-modal Turkish Sign Language Dataset and Baseline Methods. IEEE Access 8 (2020), 181340–181355. https://doi.org/10.1109/ACCESS.2020.3028072

[30] Franco Ronchetti, Facundo Quiroga, César Estrebou, Laura Lanzarini, and Alejandro Rosete. 2016. LSA64: An Argentinian Sign Language Dataset. In Proceedings of the XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016). Universidad Nacional de La Plata, La Plata, Argentina, 794–803. https://sedici.unlp.edu.ar/handle/10915/56764