Tools for the use of SignWriting as a Language Resource

Antonio F. G. Sevilla[1]**, Alberto Díaz Esteban**[2]**, José María Lahoz-Bengoechea**[3]
[1]Knowledge Engineering Institute,
Facultad de Psicología, Lateral 2, Campus de Somosaguas 28223 Pozuelo de Alarcón, Spain
[2]Department of Software Engineering and Artificial Intelligence,
Facultad de Informática, c/ Profesor José García Santesmases, 9 28040 Madrid, Spain
[3]Department of Spanish Linguistics and Literary Theory,
Facultad de Filología, edificio D, c/ Prof. Aranguren s/n, 28040 Madrid, Spain
Universidad Complutense de Madrid
[1]afgs@ucm.es, [2]albertodiaz@fdi.ucm.es, [3]jmlahoz@ucm.es

Abstract

Representation of linguistic data is an issue of utmost importance when developing language resources, but the lack of a standard written form in sign languages presents a challenge. Different notation systems exist, but only SignWriting seems to have some use in the native signer community. It is, however, a difficult system to use computationally, not based on a linear sequence of characters. We present the project "VisSE", which aims to develop tools for the effective use of SignWriting in the computer. The first of these is an application which uses computer vision to interpret SignWriting, understanding the meaning of new or existing transcriptions, or even hand-written images. Two additional tools will be able to consume the result of this recognizer: first, a textual description of the features of the transcription will make it understandable for non-signers. Second, a three-dimensional avatar will be able to reproduce the configurations and movements contained within the transcription, making it understandable for signers even if not familiar with SignWriting. Additionally, the project will result in a corpus of annotated SignWriting data which will also be of use to the computational linguistics community.

Keywords: Sign Language, SignWriting, Computer Vision, Sign Language Avatar, Textual Description

1. Introduction

One of the challenges in the study of sign language linguistics is the collection and representation of linguistic data. In computational linguistics, this problem is even more crippling, since data are the basis of any computational approach to a subject.

There is an increasing interest both in society and the scientific community in sign languages, and corpora have been created for many different sign languages and with varying schemes of annotation. However, most corpora are video-based, which is equivalent to the hypothetical case of corpora of oral languages being mostly based on audio recordings.

Recordings of real utterances, both of oral or signed languages, are difficult to process computationally, whether it is for searching or managing the data, or for linguistically analyzing it and finding its structure and meaning. Video is especially difficult, since the human visual system is highly sophisticated, and emulating its processes with artificial intelligence is not a solved problem yet.

In oral languages, writing poses a useful alternative to recordings, and is indeed (and maybe to a fault) the basis on which computational linguistics have been built. However, there does not exist an equivalent in signed languages. There is not a widely accepted written form for these languages, even less a literature or a corpus of real world linguistic data that can be exploited.

There exist some candidates for this, the most promising being SignWriting. SignWriting is a system that can act as a written form of sign language, or at least as a transcription system for it. It is iconic and very in-line with the visual nature of sign languages, so it is easy to understand and accept by native signers. The problem is that it is not as easy to use in the digital world, not being formed by linear strings of characters that can be quickly input with a keyboard and consumed by the many tools developed by the computational linguistics community.

We present an early-stage project for developing tools and resources that aim to facilitate the effective use of SignWriting in computers. With these tools, input of SignWriting can be as quick as writing it on paper, and no further processing by the user is necessary. Other tools will also use this input to generate related output, such as a textual description of the signer's actions or an animated avatar, which means that SignWriting will be useful as a digital representation of sign language even for users not familiar with it. This can help in the teaching of sign language, by facilitating the use of this language in computers, and also increase accessibility and inclusion of the Deaf community in the digital world.

In the next section, we give a brief overview of the problems of sign language notation, and quickly explain SignWriting and computer vision, the artificial intelligence tool to be used for its processing. Section 3 explicates the architecture of the project and its different components, and in section 4 some conclusions are drawn.

2. Background

Sign languages are natural languages which use the visual-gestual modality instead of the oral-acoustic one. This means that instead of performing gestures with the vocal organs, which are transmitted to the receiver via sound, sign languages utilize gestures of the body, especially of the hands, to transmit information visually.

	Stokoe Notation	HamNoSys	SignWriting
Three	$3^{\perp}$		
Bears	$[]\sqrt{C^{+}}/C_{X}^{\vee}\bullet$		
Goldilocks	$^{3}Y_{\vee}^{\circ}$		
Deep Forest	$\overline{B}_{d}\sqrt{B_{\wedge}^{\omega}}{>}$		

Table 1: Comparison of notation systems for sign languages, using words from an American Sign Language text for the story of Goldilocks and the Three Bears[2]. The systems compared are Stokoe Notation (Stokoe, 1980), HamNoSys (Hanke, 2004) and SignWriting (Sutton, 1995).

While oral languages have developed writing systems that represent the sounds (or sometimes ideas) of the language in a visual, abstract, and standard way, none such system has organically appeared for sign languages. Writing systems have many advantages, both to users of the language in helping them analyze it, and making structure explicit, and to linguists. To linguists, one advantage of writing systems of great relevance lately, and especially to us in the computational linguistics community, is the ease of computational treatment.

A number of systems have been developed for the transcription of sign language into written form (Stokoe, 1980; Herrero, 2003; Hanke, 2004). Most of them are intended for linguistic research and transcription of fine linguistic detail, and none of them seem to have seen universal use or the kind of standardization seen in the writing systems of oral languages.

This presents a challenge for the development of language resources. Systems which are alien to native informants of sign language require training for these users, and in limited time frames inevitably pose the question of whether the information transcribed with them really is what the signer intended. Additionally, we have found that computational tools for the management of the different notation systems are not very mature or wide-spread.

However, there is another proposed transcription system for sign languages: SignWriting (Sutton, 1995). SignWriting is a system developed by Valerie Sutton, a non-linguist, in 1974, designed specifically to write sign languages. There is much information on its use and practicalities on the website[1], and especially interesting is the comparison between some notation systems[2]. We reproduce a slightly modified excerpt in table 1.

We give a short introduction to SignWriting in the following, but Di Renzo et al. (2006) give an informative discussion of the use of this system in linguistic research, along with some notes on the challenges that notation systems present. More on this topic and on the differences between notation and writing systems can be found in Van der Hulst and Channon (2010).

2.1. SignWriting

As mentioned before, SignWriting is a system intended for the writing of sign languages. It is made up of symbols, many of which are highly iconic, that represent different linguistic or paralinguistic aspects. See for example Sutton (2009).

Different handshapes (such as a closed fist, an open palm, etc.) are depicted by figures like a square, or a pentagon, respectively. Conventional strokes can be added to these basic shapes to represent the thumb or the different fingers. The spatial orientation of the hand is symbolized by a black and white color code, among other possibilities. There are also icons for different locations on the body (mainly, parts of the head and face). Other symbols stand for changes in the handshape or the orientation, for different kinds of movements and trajectories, for contacts, for variations in the speed, and for facial expressions, including eyebrow intonation and other paralinguistic realizations. Finally, there are symbols that represent pauses and prosodic grouping, thus allowing to write full sentences.

All these symbols combine non-linearly in space to transcribe signs in a visually intuitive way. This is a most welcome characteristic for the Deaf community, inasmuch as they give preeminence to anything visual, and it makes it easier to learn for students of sign languages or any interested person.

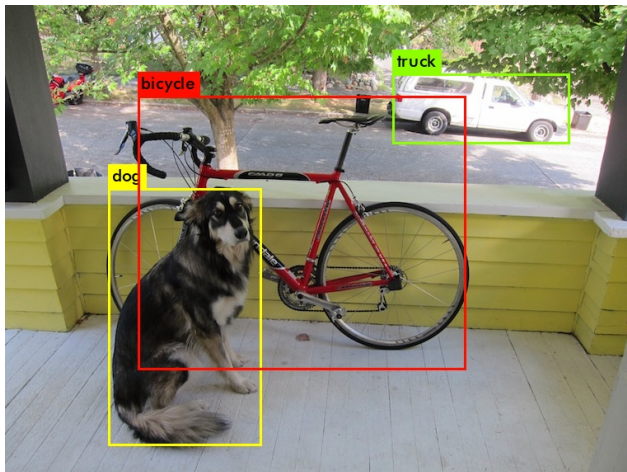Furthermore, its iconicity, together with its flexibility, al-

---

Figure 1: Object detection task, where objects in an image are located and classified (Redmon et al., 2016).

low to transcribe any newly-coined sign, making it advantageous for treating sign languages not only for daily use but also in technical, scientific, and educational environments. The fact that symbols are not interpreted linearly, but according to their position relative to other symbols, poses a challenge to the computational treatment of these bundles. It is necessary to decompose the fully transcribed signs into their components and parametrize them in linguistically relevant subunits or features.

If SignWriting annotations are created with computer tools, this information may be readily available. However, the non-linearity of SignWriting, along with the large amount of symbols that can be used, make computational input cumbersome and far slower than hand-drawing of transcriptions. Additionally, existing transcriptions, even if computer made, may not be available in their decomposed form, but rather as a plain image with no annotation. Therefore, there exists the need for tooling that can interpret images containing SignWriting transcriptions in an automatic way.

2.2. Computer Vision

Broadly speaking, computer vision is the field of artificial intelligence where meaning is to be extracted from images using automatic procedures. What this meaning is depends on the context, the available data, and the desired result. As in other fields of artificial intelligence, classification is the task of assigning a label to an image, for example the type of object found in a photograph, or the name of the person a face belongs to.

Object detection is a step beyond, in which it is to be found in an image not only what object it represents, but also where in the image it is. In the most common case, there can be many objects in an image, or none, and it is necessary to find how many there are, where, and what their labels are.

This is a difficult task, but it is very well suited to machine learning approaches, especially neural networks and deep learning. These techniques work by presenting a large amount of annotated data to the algorithm, which is able to extract from them features and patterns from which to decide the result of the procedure. Often, this means bounding

boxes: rectangles that contain the object in the image, along with labels for what the detected object is. In Fig. 1 an example of this task can be seen.

YOLO (You Only Look Once) is an algorithm for object detection that works by applying a single neural network to the full image (Redmon and Farhadi, 2018). Other algorithms work in multiple steps, for example by first performing detection of possible candidates and then classifying them, but YOLO works in a single pass, making it faster and easier to use. It works by dividing the image into regions, predicting bounding boxes and label probabilities for each region, and then collating these regions and possibilities into the final list of results. Its implementation in Darknet (Redmon, 2013 2016) is very easy to configure and utilize, while retaining precision at the state of the art.

This task of object detection is exactly what we need for understanding SignWriting transcriptions. They are formed by different symbols, placed relative to each other in a way that is meaningful and significant. By using YOLO, we can automatically find these symbols and their positions in SignWriting images, which allows us to further work with the meaning of the transcription instead of with the pixels of the image[3].

3. The VisSE project

During the authors' research in Spanish Sign Language, the problems outlined in the introduction regarding its digital treatment were patent. As students of this language as well as researchers and engineers, ideas for solutions started to come to our minds. At some point, previous expertise in image recognition, a very salient topic in sign language research, joined the knowledge of SignWriting as a useful tool for these languages, used by our educators and many in the Deaf community.

Some of the ideas for both tools and processes were combined into a single effort for which funding was requested, and granted by Indra and Fundación Universia as a grant for research on Accessible Technologies. This effort resulted in the VisSE project ("Visualizando la SignoEscritura", Spanish for "Visualizing SignWriting") aimed at developing tools for the effective use of SignWriting in computers. These tools can help with the integration of Hard of Hearing people in the digital society, and will also help accelerate sign language research by providing another methodology for its research.

A general architecture of the project can be seen in Figure 2. There, the sign in Spanish Sign Language for "teacher" is used as an example. Its transcription in SignWriting is decomposed and processed by an artificial vision algorithm, which finds the different symbols and classifies them. The labels and relative positions of the symbols are then transformed into their linguistic meaning, called here "parametrization". In the example, the usual features of sign language analysis are used, but this representation is yet to be decided, and has to follow closely the information encoded in the SignWriting transcription. The parametrization is then turned into a textual description, which allows

---

[3]When we say meaning of the transcription, we mean the codification it contains of sign language utterances, not the meaning of the represented signs in a linguistic semantics way.
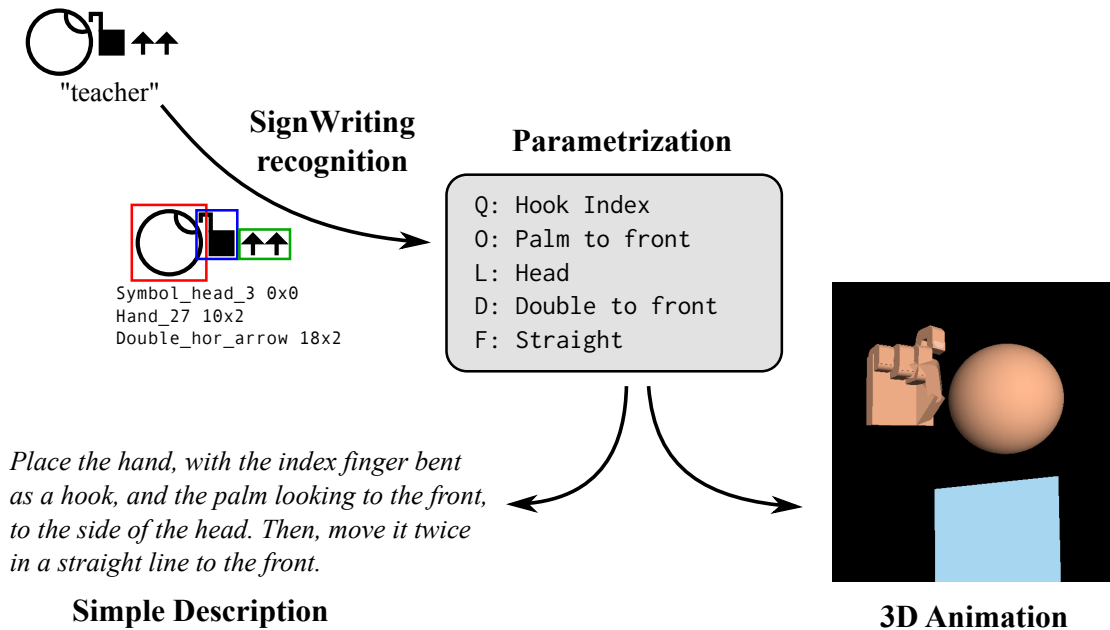
"teacher"

SignWriting recognition

Parametrization

```
Q: Hook Index
O: Palm to front
L: Head
D: Double to front
F: Straight
```

Symbol_head_3 0x0
Hand_27 10x2
Double_hor_arrow 18x2

Place the hand, with the index finger bent as a hook, and the palm looking to the front, to the side of the head. Then, move it twice in a straight line to the front.

Simple Description

3D Animation

Figure 2: Architecture of the different components of the VisSE project.

a non-signer to realize the sign, and a 3D animation which can be understood by a signer.

3.1. Corpus of SignWriting Transcriptions

While the goal of the project is to develop the tools mentioned before, which will help with the use of SignWriting in the digital world, there will be an additional language resource result. Data are of paramount importance when doing computational linguistics, and the artificial vision algorithms to be used rely on these data for their successful training and use.

Therefore, one of the products of the project will be a corpus of linguistic annotations. Entries in the corpus will be, as far as possible, input by informants who are native signers of Spanish Sign Language. For this purpose, a custom computer interface will be developed. This interface needs only be a simple front-end to the database, with roles for informants and for corpus managers, and with some tool to facilitate SignWriting input, either by a point-and-click interface or by a hand-drawing or scanner technology. Annotation, however, will not consist of grammatical information, but rather of the locations and meanings of the different symbols in the transcription.

Even if less interesting to our users, this result will probably be of use to other researchers, so it too will be publicly released. Similar to other such projects, the main object of annotation will be lexical entries, words of sign language and their realization, the main difference being that the data recorded will be in the form of SignWriting. The meaning of the annotated sign will be transcribed using an appropriate translation in Spanish.

Corpora that peruse SignWriting already exist (Forster et al., 2014), and there is also SignBank[4], a collection of tools

and resources related to SignWriting, including dictionaries for many sign languages around the world, and SignMaker, an interface for the creation of SignWriting images. While useful, the data available in the dictionaries are limited, especially for languages other than American Sign Language, and its interface is more oriented toward small-scale, manual research rather than large-scale, automated computation.

3.2. Transcription Recognizer

At first, the annotations in the corpus will have to be performed by humans, but they will immediately serve as training data for the YOLO algorithm explained in section 2 2 As annotation advances, so will increase the performance of the automatic recognition, which will be used to help annotators in their process by providing them with the prediction from the algorithm as a draft. This will accelerate data collection, which will in turn increase training effectiveness until at some point the algorithm will be able to recognize most input on its own.

The use of YOLO for recognition of SignWriting has already been successfully prototyped by students of ours (Sánchez Jiménez et al., 2019). The located and classified symbols found by the algorithm will then be transformed into the representation used in the corpus, which will include the linguistically relevant parameters (for example, it is relevant that the location is "at head level", but not whether the transcription is drawn 7 pixels to the right).

This process of finding out sign language parameters using computer vision is akin to that of automatic sign language recognition in video, which is often performed for video-based corpora. However, it is much simpler, both for the human annotator and the computer vision algorithm, since images are black and white, standardized and far less noisy. Transcriptions, being composed out of a discrete (even if
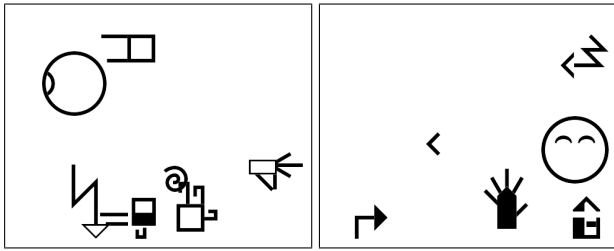
Figure 3: Automatically generated training samples for YOLO.

large) amount of symbols, present an additional advantage: training data can be immensely augmented by automatic means. An algorithm, already implemented, mixes the possible symbols to create random images which contain the data and annotations the YOLO algorithm needs to be trained. An example of this can be seen in Fig. 3. Even if these are meaningless as SignWriting transcriptions, they contain the features and patterns that the algorithm needs to learn. This will help bootstrap the system, which will make the recognizer useful to annotators earlier rather than later during the project.

3.3. Description of the Sign

Once the algorithm is able to understand transcriptions, its parametric result will be used in tools that can help integrate the Deaf community in Spain into the digital world.

The first such tool will be a generator of textual descriptions from SignWriting parametrizations. This description will explain, in a language easy to understand, the articulation and movements codified in the SignWriting symbols.

This will be a useful aid in communication between signers and non-signers. For example, it can be explained to non-Deaf people how to sign basic vocabulary, like pleasantries, or maybe important words in a particular domain (an office, a factory). This will allow non-Deaf people communicate to Deaf people basic information, useful for the daily routine or maybe an emergency, without the need to learn sign language proper. This may help broaden the employability landscape for Deaf people, increasing their inclusion in the office community and with their non-signing peers.

The use of text instead of video has some advantages. While observation of real video and images is necessary for proper understanding of the rhythm and cadence of sign language, it is often not enough for correct articulation and orientation of the hand. Non-signers are not used to looking for the visual cues in hand articulation, and may confuse the hand configurations necessary for particular signs. Spelling them out, however, can help them realize the correct finger flexing and wrist rotation, in an environment where an interpreter or teacher may not be readily available.

This also leads to a different application of this tool: education. Both Deaf and non-Deaf people can find it challenging to self-study sign language, due to the scarcity of resources and the challenge of a lack of a linear transcription system. Translating SignWriting into text can improve understanding of this system, helping both signers and students learn the representation of sign language in SignWriting symbols in a dynamic environment with immediate feedback.

3.4. Animated Execution by an Avatar

Another result of the project will be a three-dimensional animated avatar, capable of executing the signs contained in the parametric representation. SignWriting is an (almost) complete transcription of sign language, including spatial and movement-related information. This information, after its computational transformation into parameters, can be directly fed into a virtual avatar to realize the signs in three dimensions.

The advantages of the use of avatars are known in the community, and have been studied before. Kipp et al. (2011) give an informative account of different avatar technologies and the challenges in their use, and propose methodologies for testing and evaluation of the results. Bouzid and Jemni (2014) describe an approach very similar to ours in its goal, using SignWriting as the basis for generating the sign language animations. However, this process is not done automatically, but manually as in other avatar technologies.

Manual preparation of the execution of signs is a costly procedure, even if not as much as video-taping interpreters. An expert in the system as well as one in sign language are needed, and it is difficult to find both in one person. With our approach, instead of knowing the intricacies of the avatar technology, just an expert in SignWriting would be needed. It is far easier that this expert is the same as the sign language translator or author, or minimal training can be provided. SignWriting is also easier to carry around and edit, compared to systems like SIGML (Kennaway, 2004), which may be intuitive and easy for computer engineers but no so much to non-computer-savvy users.

Our system will also strive to be dynamic, not presenting a static sequence of images but rather an actual animation of the sign. While sign language generation is very complicated, it is important to note that this is not what our system needs to do. Placement and movement are already encoded in SignWriting, and our system only needs to convert the parameters into actual coordinates in three-dimensional space. The technology for this tool will be Javascript and WebGL, which are increasingly mature and seeing wide-spread adoption in the industry. Web technology is ubiquitous nowadays, and browsers present an ideal execution environment where users need not install specific libraries or software but rather use the same program they use in their everyday digital lives.

4. Conclusions

As we have seen, the goal of the VisSE project is to develop a number of tools for the use of SignWriting effectively in computers. First, the recognizer will allow SignWriting to be used as input in a comfortable way. Users will not need to search for symbols, and then drag and drop them to the canvas, nor will they have to memorize an arbitrary mapping from ASCII characters to SignWriting symbols. Hand written transcriptions or existing images will be able to be processed, making the use of SignWriting practical for continuous use. Then, the description generator and the avatar will use this input to transform SignWriting transcriptions of signs into alternative representations, which will help users understand both the meaning and the use of SignWriting.

All the developed tools will be publicly released, and the full pipeline might include software that allows a user to dynamically input SignWriting into an interface and immediately watch its realization by the avatar. The data generated in the form of the corpus can also be transformed into a dictionary, one where words are stored and indexed directly in sign language. Often, sign language resources are only accessible via oral language glosses, but the use of Sign-Writing allows sign language to be the primary language in its own dictionary.

These are all future works worthy of research and development, which will benefit the Deaf community in Spain. But the methodology and principles used are not specific to Spanish Sign Language, so we expect they will be able to be adapted to other sign languages.

Apart from the results benefiting the Deaf community, there will also be results for the language resource community. The data collected, in the form of the corpus, and the recognizer algorithm, will be released for the use of other researchers. Additionally, if this project helps SignWriting to become even more widespread and easier to use in computational contexts, this might become another powerful tool for the sign language linguistics community.

Therefore, we present this article to the community, with the goal of receiving feedback and comments during the early stages of the project so that it can inform and improve its development and its usefulness for the Computational Linguistics field.

5. Acknowledgements

6. References

Bouzid, Y. and Jemni, M. (2014). A virtual signer to interpret SignWriting. In Klaus Miesenberger, et al., editors, *Computers Helping People with Special Needs*, volume 8548, pages 458–465. Springer International Publishing.

Di Renzo, A., Lamano, L., Lucioli, T., Pennacchi, B., and Ponzo, L. (2006). Italian Sign Language (LIS): can we write it and transcribe it with SignWriting? *Proceedings of the 2nd Workshop on the Representation and processing of Sign Languages "Lexicografic matters and didactic scenarios" – LREC*, pages 11–16.

Forster, J., Schmidt, C., Koller, O., Bellgardt, M., and Ney, H. (2014). Extensions of the sign language recognition and translation corpus rwth-phoenix-weather. In *LREC*, pages 1911–1916.

Hanke, T. (2004). HamNoSys – Representing Sign Language Data in Language Resources and Language Processing Contexts. In *Proceedings of the Workshop on Representation and Processing of Sign Language, Workshop to the forth International Conference on Language Resources and Evaluation (LREC'04)*. ISSN: 17913721.

Herrero, Á. (2003). Escritura alfabética de la lengua de signos española: once lecciones. *Escritura alfabética de la lengua de signos española*, pages 1–159.

Kennaway, R. (2004). Experience with and Requirements for a Gesture Description Language for Synthetic Animation. In Gerhard Goos, et al., editors, *Gesture-Based Communication in Human-Computer Interaction*, volume 2915, pages 300–311. Springer Berlin Heidelberg, Berlin, Heidelberg.

Kipp, M., Heloir, A., and Nguyen, Q. (2011). Sign language avatars: Animation and comprehensibility. In Hannes Högni Vilhjálmsson, et al., editors, *Intelligent Virtual Agents*, volume 6895, pages 113–126. Springer Berlin Heidelberg.

Redmon, J. and Farhadi, A. (2018). Yolov3: An incremental improvement. *arXiv*.

Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788.

Redmon, J. (2013–2016). Darknet: Open source neural networks in c. http://pjreddie.com/darknet/.

Sánchez Jiménez, J. B., López Prieto, S., and Garrido Montoya, J. Á. (2019). Reconocimiento de lenguaje signoescritura mediante deep learning.

Stokoe, W. C. (1980). Sign language structure. *Annual Review of Anthropology*, 9(1):365–390.

Sutton, V. (1995). *Lessons in sign writing*. SignWriting.

Sutton, V. (2009). Signwriting: sign languages are written languages. *Center for Sutton Movement Writing, CSMW, Tech. Rep*.

Van der Hulst, H. and Channon, R. (2010). Notation systems. In Diane Brentari, editor, *Sign languages*, pages 151–172. Cambridge University Press, Cambridge.